

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) 31-05-2018		2. REPORT TYPE FINAL		3. DATES COVERED (From - To)	
4. TITLE AND SUBTITLE AIrpower: The Ethical Consequences of Autonomous Military Aviation				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) JOHN C. HEINS, Major, USAF Paper Advisor (if Any): Dr. Timothy Schultz				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval War College 686 Cushing Road Newport, RI 02841-1207				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Distribution Statement A: Approved for public release; Distribution is unlimited. Reference: DOD Directive 5230.24					
13. SUPPLEMENTARY NOTES A paper submitted to the Naval War College in partial satisfaction of the requirements of the Graduate Certificate in the Ethics of Emerging Military Technology. The contents of this paper reflect my own personal views and are not endorsed by the NWC or the Navy.					
14. ABSTRACT Simultaneous advances in remotely-piloted aircraft and artificial intelligence (AI) are converging on a likely new military capability: fully autonomous flying weapons, capable of selecting and engaging targets without direct human control. This may bring many advantages: such systems could outperform human pilots, making faster decisions and fewer errors than humans. They might also distinguish enemies from bystanders more effectively than current weapons, reducing collateral damage and civilian casualties. If these capabilities become reality, a nation might be considered irresponsible if it failed to implement them. In addition to the utilitarian benefits, potential disadvantages demand study. AI decision-making cannot always be fully explained; who could be held responsible for "Acts of Code," when AI-powered weapons make bad or indecipherable decisions? Might the mere existence of AI weaponry affect decision-makers' calculus, lowering the threshold for war? Would the humans who wield AI weapons develop an unhealthy relationship with violence? Could nations employing these weapons, in an effort to lower the cost of the war in military lives, raise the cost in civilian lives when enemies resort to terrorism? Consequentialism, deontology, and virtue ethics, along with Just War theory, provide philosophical backing for the consideration of these questions. Ultimately, and inescapably, war is a human endeavor, and those who wage it must preserve the appropriate level of human involvement. Above all, the principles of Just War must continue to supersede any technological considerations in warfare.					
15. SUBJECT TERMS Lethal autonomous weapon systems, automation, autonomy, artificial intelligence, AI, Acts of Code, ethics, consequentialism, deontology, virtue ethics, Just War, jus ad bellum, jus in bello					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES 84	19a. NAME OF RESPONSIBLE PERSON Dr. Thomas Creely
a. REPORT UNCLASSIFIED	b. ABSTRACT UNCLASSIFIED	c. THIS PAGE UNCLASSIFIED			19b. TELEPHONE NUMBER (include area code) 401-841-7542

This Page Intentionally Left Blank

NAVAL WAR COLLEGE
Newport, R.I.

The Ethical Consequences of Autonomous Military Aviation

by

John C. Heins
Major, United States Air Force

A paper submitted to the Faculty of the Naval War College in partial satisfaction of the requirements for the Graduate Certificate in the Ethics of Emerging Military Technology (EEMT) program.

The contents of this paper reflect my own personal views and are not necessarily endorsed by the Naval War College the Department of the Navy, or the Department of the Air Force.

JOHN C. HEINS, Major, USAF

31 May 2018

DR. TIMOTHY SCHULTZ
Faculty Mentor

DR. TIMOTHY SCHULTZ
Associate Dean for Electives
and Research

DR. THOMAS CREELY
EEMT Program Lead

This Page Intentionally Left Blank

Contents

Abstract.....	iv
Acknowledgements.....	v
Introduction.....	1
A Note about AI, and a Disclaimer	4
Chapter 1. The Automation Hierarchy: Machines' Long March to War	
Introduction	7
An Automation Hierarchy	8
Automation Motivations.....	9
Tier 1: Efficiency.....	10
Tier 2: Scalability	11
Tier 3: Effectiveness.....	14
Tier 4: Safety	17
Implications, and the Fifth Tier.....	21
Tier 5: Survival.....	22
Conclusion.....	25
Chapter 2. Acts of Code: When Automation Outgrows Human Liability	
Introduction	27
Automation, Liability, and the Law	28
Complexity and the Diffusion of Responsibility.....	36
Acts of Code.....	40
Conclusion.....	45
Chapter 3. Distancing and the Humanity of War	
Introduction	47
Introduction to Ethics Theories and Just War Primer	50
The "Light Side" of Distancing: Safety for Friendly Forces.....	52
The First "Dark Side" of Distancing:	
Lowering the Threshold for War	56
The Second "Dark Side" of Distancing:	
An Unhealthy Relationship with Violence	62
The Third "Dark Side" of Distancing: Terrorism	66
Conclusion.....	69
Conclusion	72
Bibliography	77

Abstract

Simultaneous advances in remotely-piloted aircraft and artificial intelligence (AI) are converging on a likely new military capability: fully autonomous flying weapons, capable of selecting and engaging targets without direct human control or intervention. The advantages this might bring are too appealing to ignore: such systems could outperform human pilots both physically and cognitively, shortening their tactical decision cycle beyond any human ability to keep pace. They could be less prone to error than humans in routine operations. They might distinguish enemies from bystanders more effectively than current weapons, reducing collateral damage and civilian casualties. If these capabilities become reality, a nation might be considered irresponsible if it failed to implement such technology to the maximum extent possible. In addition to the utilitarian benefits of such technology, its potential disadvantages demand consideration. AI decision-making cannot always be fully explained; who could be held responsible for “Acts of Code,” when AI-powered weapons make bad or indecipherable decisions? Might the mere existence of such highly-capable weaponry affect decision-makers' calculus, effectively lowering the threshold for going to war? Would the humans who wield AI weapons develop an unhealthy relationship with violence? Could nations employing these weapons, in an effort to lower the cost of the war in military lives, exact a terrible cost to civilian lives when their enemies resort to terrorism to strike back? Consequentialism, deontology, and virtue ethics, along with Just War theory, provide philosophical backing for the consideration of these questions. Ultimately, and inescapably, war is a human endeavor. Those responsible for waging it must strive to preserve the appropriate level of human involvement in war. Above all, the principles of Just War must continue to supersede any technological considerations in warfare.

Acknowledgements

No page of this project, except for this one, represents my work alone. First and foremost, I would like to acknowledge my faculty mentor for this project, Dr. Timothy Schultz. His enthusiasm, encouragement, experience, and keen eye for detail made it possible for me to transform vague, unstructured thoughts into words on paper. His steady advice as the foundations of the project shifted beneath me helped me to keep my balance and to maintain a 10,000-foot view of an effort which took on a life of its own and grew wildly beyond my expectations. I would also like to thank MIT research intern Anuj Krishnamurthy, who ventured into dusty corners of the internet to retrieve obscure facts I simply couldn't live without. Thanks also to my fellow EEMT candidates, whose knowing looks over the Eccles Library dividers gave me the reassurance I needed that I wasn't alone. Particular thanks go to research librarian Isabel Lopes, who was always eager to help, and who volunteered for the dreary task of reviewing my citations. Finally, I owe a tremendous debt of gratitude to my wife, Kate, and my daughter, Clara, who were promised much more family time this year than they received. Their love, patience, support, and Kate's extraordinarily insightful comments on my work made this paper—and this year—more academically and personally rewarding than I could possibly have dreamed.

This Page Intentionally Left Blank

Introduction

War is a human endeavor. Humans invented it in antiquity, and as their civilization evolved over the millennia, they continued to reinvent it. At first humans fought for basic resources—food, water, land. As civilization forged on, the humans’ objectives evolved: fortresses, riches, prestige, power. Later, they fought for more industrial resources like oil, ores, and minerals. Sometimes, they fought for ideologies: religion, freedom, communism, democracy. Along the way, they never stopped fighting for food, water, land, and power.

As the objects of war shifted and evolved, so did its instruments. Rocks gave way to arrows, and eventually to firearms. For thousands of years, war’s geometry was fairly stable: humans fought in two dimensions, freeing themselves of gravity just enough to fling projectiles like spears, arrows, and stones—and eventually bullets, cannonballs, and artillery rounds—as far into enemy territory as they could. Just as two-dimensional warfare was culminating in the horrors of World War I, humans learned to fight in the air. Their leap into the third dimension fundamentally shifted the character of war and strained the notions of territory, strongholds, and lines of battle. Soon they learned to fly into space, where territory and borders were meaningless. More recently, they created a virtual realm which both linked and transcended land, sea, air, and space. It made information—and the speed with which it could be acquired and acted upon—a dominant factor in warfare. Now, for the first time in this ten-thousand-year journey, humans may have to face a bizarre reality: they have advanced warfare to the point that they can no longer fight it themselves. Not without help.

Help has come in the form of automation. If there is too much information for humans to understand, they can design a machine to help them simplify it. If humans are too slow to shoot down a missile threatening their ship, they can make a weapon that will destroy the missile

automatically. If humans' reaction time is too slow to fly a jet, they can make a jet that will, at least partly, fly itself. As these technologies mature and combine, they can form something unique to the modern era: a weapon that can move, decide, and fight, all without the help of a human. Humans, the inventors and re-inventors of warfare, are on the verge of inventing ways to remove themselves from it—or, at least, to create tremendous distance between themselves and its violence.

The prospect of replacing humans with autonomous weapons immediately raises questions: Is it truly possible? Is it ethical? Both questions are vital. The first question, whether it is possible, seems to be a matter of time. Computer technology has advanced exponentially according to Moore's 1965 law, repeatedly clearing impediments to maintain its breakneck pace. For that reason, this paper makes the assumption that replacing human warfighters, and in particular combat pilots, with automated systems is possible, and will be practical in the next decade or two. This is not to say that the technologies involved are only suited to airborne combat; there are potential applications in every domain, including land, sea, space, and cyberspace. But as "drone strikes" have become the stuff of dinner table conversation, airpower provides a useful point of reference for discussing the issues which arise from the automation of war. Those issues, like the technology itself, are also not limited to airpower.

Having assumed it is possible to substitute automation for human pilots, this paper turns to the ethical implications of such a substitution. The question looms large of whether it is morally right to allow a machine to "decide" to kill a human. This paper will address that question, but there are also other, subtler dimensions of the subject which bear examination. Might machines become so competent and precise that it would be improper, and perhaps unethical, to allow a fallible human to perform a given military task? Who should be considered

at fault when one of these extraordinary machines makes a mistake with terrible consequences? Is there some vital, intangible essence of ethical warfare which might be lost if humans step too far away from its ugly violence? These are the questions at the heart of this paper.

The first chapter seeks an understanding of why humans automate tasks. To structure the discussion, it frames a hierarchy of automation—a progression, increasing in urgency, of rationales for reducing or eliminating the human role in various processes. At each tier, automation provides some sort of security, such as financial or physical security. At the speculative peak of the hierarchy, automation is the key to national security and survival. When safety and survival are on the line, failure to automate might be seen as irresponsible and negligent.

While the first chapter shows how some automation could be seen as an ethical necessity, the second chapter examines a murky corollary. Automated systems may be amazingly safe and precise, but due to the sheer complexity of their construction and of their expected tasks, they cannot be perfect. When they make a mistake—perhaps a catastrophic one—who should be held responsible? In all anticipated uses of AI on the battlefield, there will still be a human in the causal chain—perhaps many layers removed from the battlefield itself. That human might have no possible way of anticipating every outcome of the AI's actions, and in fact *no* human might have the means to do so. In some circumstances, it might not even be possible to determine why the AI made the decision it did. The chapter discusses the reasons AI mistakes are likely to be fixtures of automated combat, and suggests a term for categorizing such incidents when they occur and cannot be otherwise explained: *Acts of Code*.

The third chapter steps somewhat away from the strictly utilitarian valuation of automated warfare, and instead asks whether the psychological and physical distance created by

that type of war is problematic for any other reason. It does so in the framework of the Just War tradition, and with the help of several schools of ethics. It investigates whether access to a form of warfare less hazardous to one's own people lowers the threshold for entering a war; whether waging a war remotely, or by AI proxy, fosters an unhealthy relationship with violence; and whether prosecuting a war by automated means might increase the likelihood of retaliation in the form of terrorism, creating danger for civilians who entrust their well-being to the military. It concludes that regardless of the weapons of war, Just War determinations are inseparably human. AI weapons may have the power to distort or even corrupt national leaders' decisions, and those leaders—and the populations who choose them—have a responsibility to rise above that corrupting influence.

War is humanity's creation, and its burden. We can delegate the actions of war to machines, but we cannot transfer our responsibility. Each step along the automation hierarchy, when applied to military technology, has the potential to make it more difficult for military and government leaders to connect meaningfully with the ultimate reality: wars end human lives. For that reason alone, the prosecution of war deserves leaders' careful human attention, and their deliberate resistance to the distancing which accompanies a war waged, in part, automatically.

A NOTE ABOUT AI, AND A DISCLAIMER

The phrase "Artificial Intelligence" often conjures science-fiction images of a robot uprising. While the so-called "Singularity"—a hypothetical future point when computer intelligence will suddenly leave human intelligence far behind—may be in our future at some point, that is not the type of AI explored by this paper. AI of that sort is sometimes called "general AI," whereas AI designed to perform very specific, formerly-human tasks is termed

“narrow AI.” Narrow AI is already in widespread use, and may be found in a typical household: email spam filters, smart speakers, and self-stabilizing quadcopters all incorporate elements of AI. The title banner for this paper—styled “AIrpower” to emphasize AI’s ongoing fusion into military aviation—was created by the author in part using a Google service called the “Deep Dream Generator.” The generator repurposes AI-powered image recognition algorithms to produce new images.¹ Narrow AI is being applied to new tasks each day, and militaries and governments have taken note.

This paper uses terms such as “AI weaponry,” “AI-enabled weapons,” “AI warplanes,” “autonomous weapons,” “lethal autonomous weapon systems (LAWS),” “automated weapons,” and “killer robots” more or less interchangeably. While LAWS is the term most consistently used in policy concerning such devices, the variety of terms here is intended—besides to avoid repetitive prose—to subtly remind the reader that this is not a discussion about one particular platform, but about a revolutionary class of weapons which could take hundreds or thousands of forms.

By way of disclaimer, the author wishes to note explicitly that he is not a pilot, although his time in the Air Force has certainly imbued a healthy share of air-mindedness. Nor does he advocate that pilots be barred from the skies and relegated to ground control stations and AI tasking centers. To draw out potential issues, it was instructive to conduct a thought experiment wherein pilots of *all* aircraft were replaced by automation. As Carl von Clausewitz noted, “In the field of abstract thought the inquiring mind can never rest until it reaches an extreme.”²

Exploring this extreme highlighted cases where automation seemed most counterintuitive or

¹ Deep Dream Generator, accessed May 21, 2018, <https://www.deepdreamgenerator.com/>.

² Carl von Clausewitz, *On War*, trans. and ed. Michael Howard and Peter Paret. (Princeton: Princeton University Press, 1976), 78.

problematic. Those cases were fertile ground for considering why automation might be unwise or unethical. Finally, adopting such an extreme was helpful to reduce the influence of the author's own skepticism of automation. If it is *possible* to automate something, then a consideration of the implications of that automation is necessary, no matter how unlikely it may seem.

Chapter 1

The Automation Hierarchy: Machines' Long March to War

INTRODUCTION

Autonomous weapons have inspired a roiling ethics debate. Most of the discussion focuses on the question of whether it will be ethical to employ autonomy in a military capacity. A largely overlooked variation of that question is whether it will be ethical *not* to, once the technology matures. To explore this notion, it is useful to consider the historical reasons for automation, and then to apply the observed trends and justifications to anticipated future technology capable of autonomous function. From a utilitarian vantage point, automation offers the compelling possibility of banishing human error.

This chapter considers the possible ethical imperative to automate the human role in complex systems when technology makes it practicable. As a framework, it follows a “hierarchy of automation” which depicts, in broad categories, the reasons humans, organizations, and industries automate tasks. First, it reviews industries and sectors where automation has already supplanted human control, and the reasons for that change. There are already some industries where a return to human control would be considered unconscionable. Next, it considers modern aviation safety systems which prioritize calculation over human judgment, and how those systems fit into the hierarchy. Having established that foundation in history and technology, it will extrapolate to likely future technology and explore the possibility that concern for safety, accuracy, and combat effectiveness will compel a moral decision—not just a technical one—to remove humans, and human control, from the cockpits of aircraft.

AN AUTOMATION HIERARCHY

Figure 1 depicts an Automation Hierarchy, illustrated as a pyramid. Along the front face of the pyramid are tiers representing increasingly compelling reasons for automation: efficiency, scalability, effectiveness, safety, and survival. (“Safety” and “survival” are sometimes synonyms, but here “safety” refers to protection from accident or error, while “survival” refers to overcoming an adversarial threat.) A single technology may operate at multiple levels of the hierarchy or may migrate up the pyramid as it becomes more critical for its industry. Along the side of the pyramid are levels or types of security which adopters of automation hope to achieve. At the lower levels of the hierarchy, automation might provide financial security. At the apex, national security.

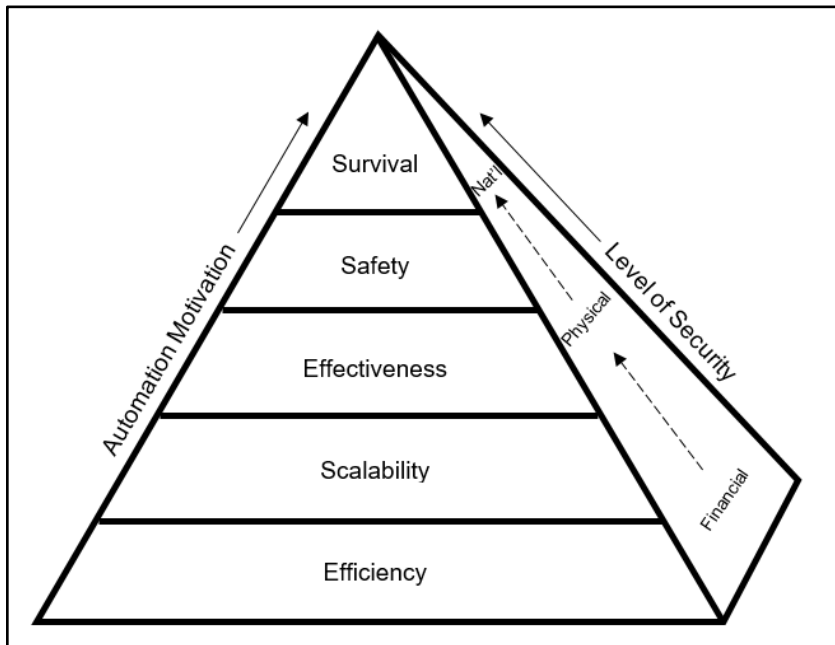


Fig. 1: The Automation Hierarchy

AUTOMATION MOTIVATIONS

For decades, automation has been closely associated with manufacturing. From the earliest days of Henry Ford's assembly line to today's highly-roboticized factories, the manufacturing industry has marched steadily toward full automation—and away from human workers. The reasons are fairly straightforward: automation can increase speed, reduce costs, and eliminate human error. Similar reasons (with different interpretations of “cost”³) will drive the military to consider such technologies for combat systems as well, as we will see.

Industries do not simply implement new technologies because they exist; more often, internal stresses or needs of industries drive the development of new technologies to meet those needs. Automobile manufacturing machines, for example, were invented to build cars better and faster.⁴ Even after such an invention, there is no guarantee that a technology will be immediately adopted; industries may resist new technologies until they can no longer deny the need for them. The military has been particularly guilty of this, such as when the British Army resisted mechanization in the interwar years. The Royal Air Force's Arthur “Bomber” Harris found himself in trouble at the UK's Army Staff College when he wrote that the army, enamored of its horse-based cavalry, would not be happy “until somebody invented tanks that ate hay and thereafter made noises like a horse.”⁵ All the same, when the need became clear, the army adopted tanks and eventually did away with its horses. To explore industries' reasons for automation more fully, we will briefly consider each tier of the hierarchy.

³ I.e., the cost in human lives rather than dollars.

⁴ Technology professor Melvin Kranzberg codified this observation with his second law of technology: “Invention is the mother of necessity.” Melvin Kranzberg, “Technology and History: ‘Kranzberg's Laws,’” *Technology and Culture* 27, no. 3 (1986): 548, <http://www.jstor.org/stable/3105385>.

⁵ Arthur Harris, *Bomber Offensive* (New York: The MacMillan Company, 1947), 24.

Tier 1: Efficiency

For some industries, automation has been implemented in the name of efficiency. For a business, “efficiency” is ultimately measured in dollars. Automation for efficiency, therefore, can provide financial security or advantage. Manufacturing and mass production are examples of this phenomenon: the need for shirts, staplers, or graham crackers has not outstripped the human capacity to produce them manually, but automation has allowed the makers to meet demand while simultaneously lowering costs.

In the public sector, where profit does not have the same primacy, reducing time and effort are still driving factors for automation. Library card catalogs, once managed meticulously by librarians to help patrons locate resources through a system of cross-references, have been almost entirely supplanted by computerized databases. The Online Community Library Center, one of the last organizations to print catalog cards, printed its last one in 2015.⁶ This transition was not undertaken to save paper; rather, it represented a more efficient way to find books. On a larger scale, the internet itself, upon which digital card catalogs rely, was initially created as a means to more quickly transmit messages among universities and government offices—an efficiency.⁷ The digitization of card catalogs and the proliferation of the internet are part of an ongoing revolution in information storage, retrieval, and sharing which has liberated knowledge from physical constraints worldwide. No longer do the answers to many queries lie in dusty volumes on high shelves; instead, they are at the fingertips of users around the world.

⁶ Online Community Library Center, “OCLC Prints Last Library Catalog Cards,” October 1, 2015, <https://www.oclc.org/en/news/releases/2015/201529dublin.html>.

⁷ “Internet History: From ARPANet to Broadband,” *The Congressional Digest* 86 no. 2 (2007): 35, <http://congressionaldigest.com/issue/network-neutrality/internet-history/>.

In each of these cases, it would still be possible, albeit inefficient, to return to “the old way”—humans performing the tasks they once did. When manufacturing or food production is done by hand, it can be considered artisanal, luxurious, or gourmet. One can still buy handmade cars and clothes, provided one is willing to pay enough. (It should also be noted that in some cases “the old way” is still more economical: the majority of smartphones are assembled by hand in Chinese factories, where inexpensive labor makes automation less appealing as a cost-saving measure. But even Foxconn, Apple’s primary assembler in China, has announced plans to automate tens of thousands of presently human jobs.⁸) People could return to using books, experts, and the post office rather than relying on the internet, but in so doing would lose tremendous benefits in speed, convenience, and cost. In general, automation undertaken in the name of efficiency is reversible, but not without losing synergy which has resulted from that automation.

Tier 2: Scalability

Some other industries, meanwhile, adopt technology by necessity, as a means to keep pace with demand or other external factors. The case of telephone switchboard operators provides an example. As recently as the mid-20th century, phone calls were completed by speaking to an operator, who would manually connect the call by plugging and unplugging cables. In the 1940s, AT&T employed 350,000 telephone operators, a high water mark.⁹ As

⁸ Nick Statt, “iPhone Manufacturer Foxconn Plans to Replace Almost Every Human Worker with Robots,” *The Verge*, December 30, 2016, <https://www.theverge.com/2016/12/30/14128870/foxconn-robots-automation-apple-iphone-china-manufacturing>.

⁹ AT&T Archives, “Her Right Place,” AT&T Tech Channel, release date July 1, 2013, <http://techchannel.att.com/play-video.cfm/2013/7/1/AT&T-Archives-Her-Right-Place>.

telephones approached ubiquity,¹⁰ this model proved insufficiently scalable. In response, telephone companies developed and implemented automatic switching equipment, reducing the number of operators at AT&T to 40,000 by 1984,¹¹ and eventually to zero.¹²

Similarly, and more recently, the banking industry has implemented electronic check clearance. The system was designed when it became apparent that existing systems would not be able to process ever-increasing quantities of paper checks. In 1994, the US Federal Reserve mandated that checks be processed electronically, eliminating a multi-day process involving the physical transportation of checks or diskettes to a clearing house.¹³ The paperless capability has been extended directly to customers in recent years: they can now deposit funds by photographing checks with their smartphones and then destroying the checks on the spot. While originally conceived to address scalability issues, these automated systems have greatly increased consumers' and business owners' ability to rapidly transfer funds, and have also reduced the opportunity to take advantage of the delay between writing a check and its clearing to commit various types of fraud. This benefit has become so integral that the Association for

¹⁰ World Bank, "U.S. Fixed Telephone Subscriptions (per 100 People)," World Bank Open Data, accessed May 17, 2018, <https://data.worldbank.org/indicator/IT.MLT.MAIN.P2?end=2016&locations=US&start=1960>.

¹¹ Sara Rimer, "Once a Friendly Fixture, a Telephone Operator Finds Herself Obsolete," *New York Times*, June 4, 1996, <https://www.nytimes.com/1996/06/04/us/once-a-friendly-fixture-a-telephone-operator-finds-herself-obsolete.html>.

¹² The Bureau of Labor Statistics still recognizes 6,000 "telephone operators" as of 2017, but the majority of these appear to be directing telephone calls within hospitals or other large businesses, or working with automated switchboards. The author was unable to locate any evidence of telephone operators currently functioning in the 1960s sense of the role.

Bureau of Labor Statistics, "Occupational Employment Statistics," United States Department of Labor, accessed May 17, 2018, <https://www.bls.gov/oes/current/oes432021.htm>.

¹³ Federal Reserve Bank of New York, "Automated Clearing Houses (ACHs)," accessed May 17, 2018, <https://www.newyorkfed.org/aboutthefed/fedpoint/fed31.html>.

Financial Professionals Electronic Payments Survey in 2013 revealed that 39% of financial professionals considered fraud control a leading benefit of electronic transactions.¹⁴

Mail-order shopping is a third industry which has leaned heavily on technology both to create growth and to respond to it. Once the idea of purchasing goods over the internet was understood and trusted by consumers, it quickly grew to overtake paper catalogs and many brick-and-mortar businesses as well. According to the US Census Bureau e-commerce has seen a 15-fold increase as a portion of total retail since 1999, from 0.6% of quarterly retail sales to over 9% in 2017.¹⁵ The popularity of online shopping forced internet retailers to innovate to prevent outages on major retail days such as “Black Friday” and “Cyber Monday.” To survive in a competitive marketplace, online retailers needed to scale to meet demand. Innovations such as redundant server farms and warehouse robots¹⁶ have made web-based commerce reliable enough to cement its place as a cornerstone of modern economic strength, and have helped the internet to grow into an integral part of daily life.

By contrast with the first group, none of these sectors could ever return to the methods they used prior to automation. Due to their growth, or the nature of the problems they were designed to solve, they would simply cease to function if they abandoned automation, and the international economy would suffer. Few would consider it “quaint” to speak to an operator for every phone call, to wait a full business day for checks to clear, or to make purchases by

¹⁴ Association of Financial Professionals, *2013 AFP Electronic Payments Survey* (Bethesda, MD, 2013), accessed May 1, 2018, 3, <https://www.afponline.org/docs/default-source/default-document-library/pub/2013-afp-electronic-payments-report-preview.pdf?sfvrsn=2>.

¹⁵ U.S. Census Bureau, “Quarterly E-Commerce Sales, 1st Quarter 2018,” accessed May 17, 2018, https://www.census.gov/retail/mrts/www/data/pdf/ec_current.pdf.

¹⁶ Sam Shead, “Amazon Now Has 45,000 Robots in its Warehouses,” *Business Insider*, January 3, 2017, <http://www.businessinsider.com/amazons-robot-army-has-grown-by-50-2017-1>.

dictating product numbers over the phone. These are industries where technology, driven by problems of scalability, has enabled tremendous leaps forward from which there is no return.

Tier 3: Effectiveness

Mass production, undertaken for efficiency (and concomitant profitability), has an additional benefit of producing extremely consistent products. The automated systems are capable of a higher degree of precision than the humans they replaced. For some tasks, that degree of precision is required in order for the task to be possible at all. One such task is precision bombing. In this case, technology was not implemented explicitly to increase speed or reduce financial cost, but to reduce human error.¹⁷ During World War II, without automated bombsights, the United Kingdom's Royal Air Force undertook a study to determine how closely bombs were falling to their intended targets during night bombings. The report found that only one in three bombs fell within five miles of its intended target.¹⁸ This inaccuracy necessitated hundreds of bombs and bombers to destroy a given target, with the undesirable additional effect of widespread devastation and loss of innocent life in the target's vicinity. In part because precision was not an option, military leaders turned instead to "area bombing" and to intangible targets such as the enemy's will to fight, which leaders thought they could drain by leveling entire cities. Ethical or tactical considerations, in varying degrees, drove the combatants to seek a better solution.

¹⁷ And, therefore, to increase "speed" of mission accomplishment and reduce the "cost" of needless human lives lost.

¹⁸ David Benussan-Butt, "Night Photographs in June–July 1941: A Statistical Analysis" (1941), 1, <https://etherwave.files.wordpress.com/2014/01/butt-report-transcription-tna-pro-air-14-12182.pdf>.

The Nazis attempted to increase their bombers' accuracy with radio beams originating in various parts of the European mainland such that they would intersect at the point when the pilot should release the bombs. The pilots were obligated to follow one beam until they crossed the other (an early case of human pilots ceding some of their autonomy to technology). Some of the later systems automated the release of the ordnance in an effort to add additional precision, removing even that task from the humans aboard.¹⁹ These systems, *Knickebein*, *X-Gerät*, and *Y-Gerät*, enabled pilots to drop their bombs in a consistent location, but did not account for weather conditions or permit numerous aircraft to attack independent targets. Additionally, they were vulnerable to compromise and countermeasure, and the British were successful in dramatically reducing the Germans' effectiveness through what may be considered history's first example of electronic warfare.²⁰

The United States, meanwhile, turned to the famous Norden bombsight. While it never achieved the pinpoint accuracy in combat which it demonstrated in peacetime testing,²¹ it provided a means to attack a specific point on the land with much greater accuracy than the 5-mile radius identified in the British study, and once paired with the Minneapolis-Honeywell C-1 autopilot in 1943, achieved significant battlefield success. In one case in March 1943, greater than three-quarters of bombs struck within 1,000 feet of their targets.²² A key element of the Norden bombsight's functionality was its ability to actively fly the plane during a bomb run. Because human pilots could not keep the plane stable enough for the bombsight to operate

¹⁹ Richard Overy, *The Bombing War: Europe 1939-1945* (London: Penguin Press, 2013): 76-77.

²⁰ David Gates, *Sky Wars* (London: Reaktion Books, 2003): 44.

²¹ John Guttman, "Norden M-1 Bombsight," *Military History* 25, no. 5 (2008): 25, <https://search-proquest-com.usnwc.idm.oclc.org/docview/212664512>.

²² Raymond P. O'Mara, "The Socio-Technical Construction of Precision Bombing: A Study of Shared Control and Cognition by Humans, Machines, and Doctrine During World War II" (PhD diss., Massachusetts Institute of Technology, 2011), 163, <http://hdl.handle.net/1721.1/67754>.

correctly, the bombsight was tied into the plane's autopilot. For at least 50 seconds in training,²³ and likely longer in combat—an eternity in flak-filled enemy skies—the pilot would cede control to the bombsight, which would fly the plane until it reached the calculated place and time to release its munitions. Even more than the German system, the Norden bombsight required the pilot to relinquish control to an automated system which could, for those long moments, do his job better than he could. Not all pilots accepted this shift in power dynamics easily; one bomb wing commander had to threaten pilots with courts-martial in order to convince them to cede control to the bombsight.²⁴

Since the days of the Norden bombsight, the ability to target with precision has only improved. The world watched in astonishment as the U.S. showcased video footage of pinpoint strikes during the first Gulf War, and today's laser- and GPS-guided munitions are a quarter-century further advanced. Modern weapons no longer require pilots to hand over control in such dramatic fashion, but they are still automating tasks which used to be done by humans, with far greater success. To go back to manual targeting and bombing would be beyond quaint, beyond inefficient, and beyond impractical: it would be criminally negligent in light of the technology available that so effectively minimizes collateral damage. In this case, there is an ethical imperative to automate the task. As will be further explored in Chapter 3, failure to use such automation would jeopardize a force's ability to wage a just war.

²³ O'Mara, 161.

²⁴ Gerald Astor, *The Mighty Eighth* (Dutton Adult, 1997), 164. Quoted in O'Mara, 163-4.

Tier 4: Safety

Another common application for automation is to compensate for human error with technological safeguards. These systems seek to improve humans' physical security by taking some degree of control away from them. Newer vehicles, on the march toward self-driving cars, will notify the driver if he or she strays from the lane, or if the car detects an object in a blind spot. Some vehicles will automatically brake to prevent a collision if the driver fails to do so. Regulators could eventually make such features mandatory, as they have done with reverse cameras,²⁵ in the name of reducing fatalities on the road.

Similar technologies exist for rail systems; in response to a fatal incident in 2008, legislators mandated that automatic braking for trains be installed by the end of 2018. In 2017, an Amtrak train derailed, killing three aboard. The mandatory system, known as Positive Train Control (PTC), had been installed but not yet been activated because it was still being tested.²⁶ The National Transportation Safety Board (NTSB) reported that the train's engineer had failed to see speed-limit signs until it was too late—an error an automated system would be unlikely to make.²⁷ It is likely that if PTC had been fully operational, those three lives would have been saved.

Aircraft have safety systems of their own to compensate for human error. In 2009, U.S. Airways flight 1549 lost both engines to bird strikes, and the pilot, Chesley "Sully" Sullenberger,

²⁵ National Highway Traffic Safety Administration, "NHTSA Announces Final Rule Requiring Rear Visibility Technology," March 31, 2014, <https://www.nhtsa.gov/press-releases/nhtsa-announces-final-rule-requiring-rear-visibility-technology>.

²⁶ Bart Jansen, "Amtrak Train Derailed on Tracks that had Automatic-braking Technology—But it was Still Being Tested," *USA Today*, December 19, 2017, <https://www.usatoday.com/story/news/2017/12/19/speeding-amtrak-train-derailed-track-without-automatic-braking/964483001/>.

²⁷ Mike Lindblom and David Gutman, "NTSB Report: Amtrak Engineer Missed Speed-limit Signs Before Train Crashed South of Tacoma," *Seattle Times*, January 25, 2018, <https://www.seattletimes.com/seattle-news/transportation/ntsb-report-amtrak-engineer-missed-speed-limit-sign-before-the-train-crashed-on-a-curve-south-of-tacoma/>.

was famously forced to land on the Hudson River. Following the engine losses, Sullenberger immediately activated the plane's Auxiliary Power Unit, which enabled the plane's fly-by-wire computer system to prevent excessive control inputs which might cause the crippled plane to exceed flight envelope parameters and stall. Thus, one of the foremost heroic achievements of human pilots was enabled, or at least constrained, by technology which helped him avoid some errors. It is noteworthy that the NTSB investigation determined that if Sullenberger had taken action instantly, rather than using 35 seconds to observe, orient, and decide, he could have landed at either of two nearby runways rather than on the river.²⁸ Humans are of course not capable of making such near-instantaneous decisions, but a computer might be.

Many aircraft will warn the pilot upon detection of an imminent collision with terrain, and pilots flying in conditions with poor visibility are reliant on these and countless other sensors and instruments to keep their planes upright and aloft. Some systems take it a step further, however. A feature known as the Automatic Ground Collision Avoidance System (Auto-GCAS) has been incorporated into U.S. Air Force F-16 fighter jets. It will detect an imminent collision with the ground and, after attempting to warn the unresponsive pilot, take independent action to prevent a crash. The system has already saved at least four lives, according to public reporting.²⁹ Those four people are elite, highly-trained pilots whose ability to control a high-speed, high-

²⁸ National Transportation Safety Board, *Loss of Thrust in Both Engines After Encountering a Flock of Birds and Subsequent Ditching on the Hudson River, US Airways Flight 1549, Airbus A320-214, N106US, Weehawken, New Jersey, January 15, 2009*, May 4, 2010, <https://www.nts.gov/investigations/AccidentReports/Reports/AAR1003.pdf>.

²⁹ Guy Norris, "Auto-GCAS Saves Unconscious F-16 Pilot—Declassified USAF Footage," *Aviation Week*, September 13, 2016, <http://aviationweek.com/air-combat-safety/auto-gcas-saves-unconscious-f-16-pilot-declassified-usaf-footage>.

performance, extraordinarily complex system is essentially unrivaled—and yet, without the intervention of a superior³⁰ form of control, they would have died.

While there are no humans aboard remotely piloted aircraft (RPAs), avoiding crashes is still of primary concern for their operators, and still constitutes “safety of flight.” Because of their distant link to the pilot, RPAs must be capable of some independent maneuvers. For example, they may need to take specific actions if they become disconnected from their operators—these are known as “lost link” procedures. At present, these must be pre-determined and pre-programmed, and adjusted throughout the course of a mission so that the aircraft does not, for example, land in hostile territory because of a lost link.³¹ In addition, because of the inherent delay in receiving control signals transmitted from great distances (as much as 2-3 seconds for a satellite link),³² some intricate maneuvering must be handled automatically, rather than relying on signals from the pilot. For some airframes, such as the MQ-4 Global Hawk, this includes takeoffs and landings: rather than manually piloting the landing, for example, the remote pilot commands the aircraft to land and provides it with appropriate parameters, and the aircraft performs the maneuver without further direction. In 2016, the Air Force announced that

³⁰ Superior for this circumstance, as designed.

³¹ Houston Cantwell and Alfred Rosales, “RPA Lost Link: What do we do now?,” *Combat Edge*; Langley AFB Vol. 26, Iss. 1, (Jun/Aug 2017): 14.

³² Mustapha Mouloua et al., “Ergonomics of UAV/UCAV Mission Success: Considerations for Data Link, Control, and Display Issues,” *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting*, Minneapolis/St. Paul, MN, 2001 (Santa Monica: Human Factors and Ergonomics Society, 2001), 144, https://www.researchgate.net/profile/Peter_Hancock2/publication/238075274_Ergonomics_of_UAVUCAV_Mission_Success_Considerations_for_Data_Link_Control_and_Display_Issues/links/00b7d51c8604b9173a000000/Ergonomics-of-UAV-UCAV-Mission-Success-Considerations-for-Data-Link-Control-and-Display-Issues.pdf; Rob Blackhurst, “The Air Force Men Who Fly Drones in Afghanistan by Remote Control,” *The Telegraph*, September 24, 2012, <https://www.telegraph.co.uk/news/uknews/defence/9552547/The-air-force-men-who-fly-drones-in-Afghanistan-by-remote-control.html>; Alan Hobbs, “Human Factors of Remotely Piloted Aircraft Systems: Lessons from Incident Reports,” National Aeronautics and Space Administration, last modified February 10, 2017, <https://www.nasa.gov/mediacast/human-factors-of-remotely-piloted-aircraft-systems-lessons-from-incident-reports>.

the MQ-9 Reaper would join the MQ-4 Global Hawk in performing automated takeoffs and landings.³³

Another aviation task where automation has shown promise is the aircraft carrier landing. This life-and-death test of a pilot's skill leaves so little room for error that engineers have sought technological means to ease some of the burden. Legacy airframes have been fitted with a system known as MAGIC CARPET,³⁴ which reduces the number of necessary control inputs during a landing from 300 to an average of 10-20.³⁵ Similarly, a system called Delta Flight Path has been included in the U.S. Navy's variant of the F-35 Lightning II fighter jet to assist pilots with carrier landings. In early tests, it was so accurate that the planes were literally wearing holes in the test runway by striking the exact same spot each time. Engineers were forced to program a small variance into the system—an intentional error—to reduce the expected wear on the decks of aircraft carriers from the extraordinary consistency of the system.³⁶

In all of these examples, technological solutions have been introduced to improve upon human limits, and to remove or mitigate danger and uncertainty caused by human error. While at present the technologies remain in a supporting role, it is not difficult to imagine steadily-improving technology accruing more and more responsibility until the pilot is only responsible for mission-related input to the system. Even there, the pilot's role may be in jeopardy.

³³ James Drew, "USAF to Automate MQ-9 Takeoffs and Landings," *Flight Global*, May 4, 2016, <https://www.flightglobal.com/news/articles/usaf-to-automate-mq-9-takeoffs-and-landings-424975/>.

³⁴ Maritime Augmented Guidance with Integrated Controls for Carrier Approach and Recovery Precision Enabling Technologies

³⁵ Eric Adams, "New Navy Tech Makes it Easy to Land on a Carrier. Yes, Easy," *Wired*, August 2, 2016, <https://www.wired.com/2016/08/new-navy-tech-makes-landing-aircraft-carrier-breeze/>.

³⁶ *Military.com*, "Navy F-35C Landed So Precisely, it Tore up a Runway," August 18, 2016, <https://www.military.com/dodbuzz/2016/08/18/navy-f-35c-landed-so-precisely-it-tore-up-a-runway>.

IMPLICATIONS, AND THE FIFTH TIER

If all military pilots were replaced by highly capable artificial intelligence (AI), what ethical problems would arise from a socio-technical perspective? Is there intrinsic value in today's human-machine teaming which would be lost if the machine performed all the tasks? The change might be judged in terms of lives saved: if AI flew more safely, i.e., crashed less frequently, then in the long term it would be detrimental to allow a human to fly a plane with human passengers, because their likelihood of dying would be higher—however marginally. Furthermore, an autonomous aircraft flying in defense of the nation could be more effective in some missions—and perhaps eventually all missions—than one with a human pilot.³⁷ Finally, autonomous offensive capabilities could be more potent, leading to the destruction of more enemy targets with less friendly exposure. From a purely utilitarian perspective, autonomous military aircraft would seem to have many advantages. In that respect, automated military aviation would have benefits similar to factories: humans *could* do the task, but why *should* they, when a machine can do it more efficiently, effectively, and safely?

There are many types of missions where the human element seems to be indispensable. As an example, consider the task of assessing an off-course airliner. The human pilot of an interceptor aircraft could look through the airliner's windows to attempt to determine whether the plane was experiencing a malfunction, or the pilots had been incapacitated, or the plane had been hijacked. For a human, this is an easy determination (relatively speaking). An autonomous aircraft would need to be extraordinarily sophisticated to perform the same task—but the technology to do so is not inconceivable. Additionally, consider the possibility that airliners, out

³⁷ Particularly when flying against a similarly automated foe, as we will explore.

of a concern for safety, were fully automated. There would be no pilots to be incapacitated, and any hijacking would have to be done remotely, with no visible clues.

Tier 5: Survival

As military technologies develop, the military's employment of automation and increasingly autonomous systems may come to resemble other industries where a return to human control would be impossible (such as telephone switchboards). For most industries, this would represent Tier 3, Effectiveness. Recall that some technologies may fall on multiple levels of the hierarchy depending on the context of their application. In the military's case, effectiveness can be inextricably linked to tactical relevance, and tactical relevance to survival. We may soon discover that humans simply cannot perform the tasks necessary to win a war because those tasks will have grown in complexity and velocity to such a degree that the choice is between automation and failure.

Air Force Colonel John Boyd introduced the concept of the "OODA loop" in the 1960s to describe a combat pilot's decision-making cycle. The acronym stands for observe, orient, decide, and act, at which point the loop begins again. His theory continues that whichever pilot has the "smaller," or faster, OODA loop can get "inside" his or her adversary's decision cycle and gain an advantage.³⁸ While decision-making has historically been the human's domain, speed is decidedly the computer's. An automated pilot which can make faster decisions than a human pilot will therefore have a smaller OODA loop—and increase its odds of victory.

³⁸ John Boyd, "The Essence of Winning and Losing," The Internet Archive, accessed May 17, 2018, <https://web.archive.org/web/20110324054054/http://www.danford.net/boyd/essence.htm>.

This is problematic for advocates of human decision-making, who would be more comfortable entrusting a human with the many decisions a tactical situation demands. However, this aversion to automation could be fatal. If the U.S. were to institute a policy that only a human could decide whether to fire upon another aircraft, for example, an adversary could take advantage of the delay caused by that requirement by building an autonomous system with no such compunction. The adversary's OODA loop would be smaller, and their aircraft would have a tactical advantage. To avoid defeat, the United States would have no choice but to automate its lethal decisions as well. This is scarcely hypothetical: A Russian senator has already indicated his country's intent to build autonomous combat aircraft,³⁹ and Vladimir Putin himself said that whoever masters AI "will become the ruler of the world."⁴⁰

The "observe" and "orient" parts of the OODA loop are growing more difficult every day, at least for humans. A combat pilot's world is one of information assault. The cramped cockpit, and lately even the pilot's helmet, are crammed full of data displays in the hopes of enabling the best-possible decision. The pilot's ability to absorb all of this information while also flying the aircraft is, of course, limited. Display screens have proliferated in RPA control stations in part because they are unconstrained by the space limitations imposed by an aircraft cockpit. As a result, RPA pilots have as many as nine screens to monitor, which NASA Engineering and Safety Center's Dr. Alan Hobbs concluded is far too much information to process, and is more likely to cause error than prevent it.⁴¹

³⁹ TASS Russian News Agency, "Artificial Intelligence to Replace Pilot in Aircraft Cockpit – Russian Senator," November 1, 2017, <http://tass.com/defense/973707>.

⁴⁰ David Meyer, "Vladimir Putin Says Whoever Leads in Artificial Intelligence Will Rule the World," *Fortune*, September 4, 2017, <http://fortune.com/2017/09/04/ai-artificial-intelligence-putin-rule-world/>.

⁴¹ Hobbs.

Interpreting a large volume of data rapidly is not a human's strong suit, but it is a computer's. The recent age of AI, "big data" processing, and "deep learning" algorithms has produced systems well-suited for this sort of task. As the number of sensors and data feeds on a combat aircraft grows, the task of managing a cockpit and making decisions based on available information is becoming one only a computer could manage effectively. The OODA loop of observe-orient-decide-act simply becomes input-compute-output for AI, without slowing down to wait for a human to choose a course of action.

In a war with a peer adversary, it would be reckless to limit decision-making for the sole purpose of keeping a human involved, particularly if the decisions produced were equivalent or worse than the "decisions" the AI could have produced. Once the technology sufficiently matures and is in the hands of adversaries, there will be no going back; it would be like unplugging the U.S. internet and expecting the economy to continue uninterrupted—except in this case lives and national security would be on the line.

Finally, in addition to conferring a combat advantage in the presence of an adversary OODA loop, there is the strong possibility that an AI-controlled combat aircraft will simply be more effective at routine missions. Whether it be more reliable carrier landings, more precise strikes, or another capability we currently lack, AI pilots could deliver effects which today would be considered science fiction. Imagine an autonomous stealth craft which could loiter high over an adversary city for days or weeks. It could use high-powered optics and sensors to perform cell phone geolocation, gait analysis, facial recognition, and other means of identification. Once it had located one of its many vetted and approved targets, it could dispatch an insect-size drone to pursue or observe him or her, or it could engage immediately and with surgical precision. A

human controller would add no additional tactical capability, but might cause enough delay to withhold this and countless other unimagined capabilities from a future arsenal.

CONCLUSION

If—and likely “when”—computers become safer pilots than humans, we humans will have a decision to make, and we could make it using simple math. Commercial aviation, for example, has become amazingly safe compared to its early decades. The odds of experiencing a plane crash on one of the world’s major airlines are around 1 in 3.4 million. While those odds are good, approximately 60% of the crashes which do occur are attributed to human error.⁴² If automation could eliminate those crashes, the odds of a plane crash could fall to 1 in 5.4 million.⁴³ If automated pilots assured statistically safer flying, one more plane crash attributed to human error could be considered criminal negligence on the part of the airline which failed to modernize.⁴⁴

The same logic would apply to military missions. If full automation increased the odds of a positive outcome, it could be seen as dereliction of duty to induce risk by subjecting the mission to the vagaries of human judgment and error. The pilot’s very presence could become detrimental to mission success. From a purely utilitarian perspective, it seems that removing the pilot and handing the reins over to an artificial intelligence may be an ethical imperative. This line of reasoning will doubtless make many observers uncomfortable. They may contend that

⁴² “Causes of Fatal Accidents by Decade,” PlaneCrashInfo.com, accessed May 17, 2018, <http://www.planecrashinfo.com/cause.htm>.

⁴³ This assumes that a human’s presence is never needed to recover from a malfunction which would otherwise cause a crash. This is not a safe assumption today, but more sophisticated automation systems could compensate for recoverable faults in the same way a human pilot would.

⁴⁴ Of course, automated systems themselves could malfunction, in the same vein as human pilots who make mistakes; this is the subject of Chapter 2.

there are human elements at play, particularly in the “decide” portion of the OODA loop. Even if a computer *could* decide, does that mean that it *should* decide? These questions will be explored in more detail in Chapter 3.

Chapter 2

Acts of Code: When Automation Outgrows Human Liability

INTRODUCTION

The previous chapter forecasted the existence of systems based on artificial intelligence (AI) which could undertake complex tasks, like aerial combat, with fewer errors than in human-controlled systems. No system, however, is perfect, and even the most sophisticated systems (some might say *especially* the most sophisticated systems) are bound to behave undesirably from time to time. When such failures occur, the question looms large of who should be held responsible. Consider a hypothetical case of an automated combat aircraft which destroys a hospital instead of a weapons depot for unclear reasons, resulting in hundreds of casualties. Who is responsible? The person who launched the mission? The person who approved the mission? The person who authorized the purchase of the system? The government contractor who produced it? Or, perhaps, the programmer(s) who wrote the code?

The answer is bound to be some variation of “it depends.” The countless ways in which a mission could go awry would lead to countless root cause analyses which could assign blame to any number of entities. But it is possible that the analysis could show that every person in the design, acquisition, and decision chains did his or her due diligence and had no reasonable means to foresee the error. In that circumstance, might it be that *no one* is responsible for the undesirable outcome? Scholars have asserted that humans are “innate retributivists,” meaning that we want to punish someone when a person is harmed.⁴⁵ How will we satisfy that urge if such seemingly faultless incidents become commonplace?

⁴⁵ John Danaher, “Robots, Law and the Retribution Gap,” *Ethics and Information Technology* 18, no. 4 (2015): 299.

This chapter delves into the issues surrounding automated systems and liability. It begins by discussing systems already in use around the world, as well as some efforts underway to tackle legal issues associated with near-future technology. It then discusses some circumstances where individual liability has been effectively eliminated or diffused due to the complexity of the situation or causal chain. Finally, it proposes a label for unforeseeable and unpreventable events when they are perpetrated by highly complex systems.

AUTOMATION, LIABILITY, AND THE LAW

The issue of liability in an increasingly automated world is not a new one. A 1961 column in the *Wall Street Journal* bemoaned that “[C]omputers cannot exercise the same judgement as the phlegmatic bookkeepers they replaced.” It wondered who might be responsible for financial errors resulting from automated banking, and whether an explosion resulting from a highly automated chemical plant should be considered the fault of the chemical company or of the computer company. The column does not provide easy answers, although it does cite a joint committee of the American Bar Association and the American Law Institute which suggested that “companies should have attorneys examine computer systems to assess possible legal risks.”⁴⁶ It seems improbable that the committee foresaw the degree to which computers and automated systems would proliferate and increase in complexity in subsequent years, or the consequent difficulty of effectively implementing that recommendation.

The practice of law as it relates to automation and automated processes is, 57 years later, still very immature. However, efforts are afoot in some quarters to prepare for the next evolution of legal challenges to arise from technology. Automated systems, as they currently exist, cannot

⁴⁶ *Wall Street Journal*, “Automation and Liability,” December 12, 1961, 14.

be held directly accountable for their actions.⁴⁷ Some experts have proposed that autonomous systems be enrolled in a liability fund which would enable them to pay penalties for damages, but also acknowledge that “ascribing full rights and duties to these entities is not viable at present.”⁴⁸ While no court has yet found an automated system directly liable for a crime, the concept of machines’ legal immunity may be showing its first cracks. In response to perceived abuses by the National Security Agency using automated surveillance systems, the U.S. Court for the District of Columbia ruled such surveillance unconstitutional. It determined that it was a violation of citizens’ Fourth Amendment protections from unreasonable search and seizure—even if a human never looked at the collected data, which had been the Agency’s main defense.⁴⁹ The court’s decision may have far-reaching consequences: it provides a precedent for future decisions that automated systems can violate individuals’ rights on their own, something which previously only other humans could do.

In addition, the judicial system has already begun to contend with the deadly consequences of some autonomous systems. The U.S. Occupational Safety and Health Administration lists 28 robot-related workplace fatalities dating as far back as 1987.⁵⁰ Despite being robot-related deaths, in most cases, the event has been ascribed to human error. In such historical cases, the outcome was normally a procedural or technical fix; some cases additionally resulted in charges, litigation, and fault-finding. The U.S. is not alone in contending with these

⁴⁷ Tomoko Nambu, "Legal Regulations and Public Policies for Next-Generation Robots in Japan," *AI & Society* 31, no. 4 (2016): 485.

⁴⁸ Hironori Matsuzaki and Gesa Lindemann, "The Autonomy-Safety-Paradox of Service Robotics in Europe and Japan: A Comparative Analysis," *AI & Society* 31, no. 4 (2016): 508.

⁴⁹ Meg Leta Jones, "The Ironies of Automation Law: Tying Policy Knots with Fair Automation Practices Principles," *Vanderbilt Journal of Entertainment and Technology Law* 18, no. 1 (2015): 94-95.

⁵⁰ Accident Database, Occupational Safety and Health Administration, queried January 30, 2018, https://www.osha.gov/pls/imis/AccidentSearch.search?acc_keyword=%22Robot%22&keyword_list=on

problems; the summer of 2015 was particularly eventful worldwide. In July, a worker at a German Volkswagen plant was killed when a robot he was working on was inadvertently activated and crushed him. It was reported as “human error,” and the last available information indicated that authorities were considering charges against management.⁵¹ Around the same time, a woman was killed by machinery at a factory in Michigan when a robot entered the wrong area of a factory and trapped and crushed a woman with a swing arm, killing her. The company was fined just \$7,000 for negligence.⁵² The next month, a robot killed a worker in India, either by electrocution or by “piercing,” depending on the account. “The company management and the contractor [were arrested] on charges of causing death due to negligence,” according to the local Assistant Commissioner of Police.⁵³

In each of these cases, human management was either alleged to be responsible or was found to be, with corresponding punitive action. It seems possible that this trend might continue as technology evolves, so that even as humans with agency occupy ever more distant positions from the incidents themselves, the chain of responsibility will ultimately lead back to them. But identifying who the human with agency is may become increasingly difficult as the technology evolves. This is a topic of active discussion in Japan.

Japanese experts are proactively grappling with the legal ramifications of so-called “Next Generation Robots” (NGRs), or helper robots. Because of Japan’s aging population and low birth rate, the country anticipates relying upon NGRs to assist some elderly citizens in the near future.

⁵¹ Eliana Dockterman, “Robot Kills Man at Volkswagen Plant,” *TIME Magazine*, January 2, 2015, <http://time.com/3944181/robot-kills-man-volkswagen-plant/>.

⁵² *WOOD TV*, “Fine issued in worker’s death at Ionia plant,” February 2, 2016, <http://woodtv.com/2016/02/02/fine-issued-in-workers-death-at-ionia-plant/>.

⁵³ *International Business Times India*, “Robot Kills Man at Gurgaon Factory,” August 13, 2015, <http://www.ibtimes.co.in/robot-kills-man-gurgaon-factory-642723>.

This application of robotics differs significantly from industrial use, where robots are ordinarily designed to avoid touching people or to operate separately from them.⁵⁴ Because robots cannot be directly held responsible, liability for hypothetical NGR-caused incidents could be assigned to the manufacturer, the custodian, or the operator.⁵⁵ In order to escape liability, manufacturers may have to demonstrate that it was impossible to predict the outcome—a challenging negative to prove.⁵⁶ Custodians, meanwhile, may be required to take responsibility for the actions of their NGRs, much like pet owners are. Pet owners in Japan can currently escape liability by showing that they “managed the animal with reasonable care according to its kind and nature,”⁵⁷ something which may not be easily definable for NGRs. Finally, operators (distinguished from custodians by their direct control of the NGR) may be liable if it could be proven that they caused the incident through negligence or lack of due care.⁵⁸ In the case of autonomous or semi-autonomous NGRs, the distinction between custodian and operator may become difficult to discern, because an NGR may undertake a long series of independent actions in response to a single command. While the International Organization for Standardization (ISO) and other organizations have begun to publish standards to by which due diligence might be assessed, including the ISO’s “Safety Requirements for Personal Care Robots,”⁵⁹ liability in these cases is still determined on a case-by-case basis in Japan as elsewhere.⁶⁰ To address these and related concerns, the Japanese government has created a “Robot Revolution Realization Council,” with

⁵⁴ Nambu, 484

⁵⁵ Nambu, 486.

⁵⁶ Nambu, 487.

⁵⁷ Nambu, 487.

⁵⁸ Nambu, 488.

⁵⁹ International Organization for Standardization, “ISO 13482:2014: Robots and Robotic Devices – Safety Requirements for Personal Care Robots,” accessed May 1, 2018, <https://www.iso.org/standard/53820.html>.

⁶⁰ Nambu, 497

the goal of solving these problems in time to make use of the technology as it becomes available.⁶¹

Japanese NGRs are not the only locus of thought for governing autonomous technologies. Another active sector is the automotive industry, which is grappling with the ramifications of self-driving cars—which are a reality today, albeit still in limited fashion. A car which maintains distance from the car in front of it and stays in its lane automatically is, technically speaking, driving itself. The driver’s functions of providing mechanical input to the system (accelerating, braking, and steering) have all been automated—that is, until the passenger needs the car to depart from the established course. But the more commonly understood meaning of “self-driving” can be taken from guidance published in 2014 by SAE International (formerly the Society for Automotive Engineers) which defines an automated driving system as one which, at a minimum, can take responsibility for “all aspects of the dynamic driving task with the expectation that the human driver will respond appropriately to a request to intervene.”⁶² Similarly, the United Kingdom Department of Transport defines fully-automated cars as those “designed to be capable of safely completing journeys without the need for a driver in all normally encountered traffic, road and weather conditions.”⁶³ The Tesla Model S, according to its manufacturer, ships with all the sensors necessary for “full self-driving capabilities,” although the company does not plan to make such capabilities available until it has acquired sufficient

⁶¹ Nambu., 496

⁶² SAE International, “Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems,” accessed May 1, 2018, https://www.sae.org/standards/content/j3016_201401/.

⁶³ United Kingdom Department of Transport, *The Pathway to Driverless Cars: Summary report and action plan*, (London, England, 2015), 18.

driving data and until relevant governing bodies have established appropriate policy.⁶⁴ The establishment of those policies will require consideration of some unprecedented legal questions.

Self-driving cars will be entrusted with choices that formerly only a human could make, including decisions where death may result, either for an occupant of the vehicle or someone outside of it. To further the discussion, the Massachusetts Institute of Technology created a project called “Moral Machine,” which seeks a “clearer understanding of how humans make such choices, but also a clearer understanding of how humans perceive machine intelligence making such choices.”⁶⁵ In pursuit of that goal, the Moral Machine web site presents scenarios and asks visitors to play the role of a self-driving car. The user, acting as the car, is forced to choose between “killing” the vehicle’s occupants or different groups of pedestrians—for example, the visitor might be asked to choose between killing four adult passengers or three children crossing the road. The project hopes not only to gather relevant data about the ethics of automation, but also to broaden discourse on the subject in order to positively influence future policy and design.

When a self-driving car causes an accident, whether through malfunction or an unwinnable scenario like the one described above, liability may be ascribed in one of two broad ways: to the driver, or to the manufacturer. This is similar to the NGR debate in Japan, with the driver serving as both custodian and operator. The driver, whether wittingly or not, set the car on its ill-fated course, even if he or she was not directly responsible for the fateful maneuver. Had he or she not gotten in the car, there would have been no incident, and so he or she could be

⁶⁴ Jordan Golson and Dieter Bohn, “All new Tesla cars now have hardware for ‘full self-driving capabilities’,” *The Verge*, October 19, 2016, <https://www.theverge.com/2016/10/19/13340938/tesla-autopilot-update-model-3-elon-musk-update>.

⁶⁵ Massachusetts Institute of Technology, “Moral Machine,” last accessed February 11, 2018, <http://moralmachine.mit.edu>.

found liable. The manufacturer, on the other hand, might be found liable for problems with the vehicle's design or programming—manufacturing “defects,” in essence.⁶⁶ In the latter case, the question must arise of what constitutes due diligence for the manufacturer.

Under the British definition of a self-driving car, the vehicle (and therefore the manufacturer) must account for “all normally encountered traffic, road and weather conditions,” yet does not define “normally encountered” conditions. A footnote to the British definition suggests that the car's criteria not to proceed (extreme weather, for example) might be comparable to a human driver's criteria for making the same decision.⁶⁷ But if such a car were to fail to recognize a snow-covered stop sign and, as a result, strike a pedestrian, who would be liable? The occupant, for embarking on the drive? Or the manufacturer, for failing to account for that road condition?

European legal scholars Sabine Gless, Emily Silverman, and Thomas Weigend argue that the bulk of liability should be shifted from the operator, where it currently lies, to the manufacturer for the simple reason that the operator will no longer be in a position to reliably foresee and prevent undesirable system behaviors.⁶⁸ Melinda Lohmann, a German legal scholar, agrees that manufacturers will experience an elevated level of scrutiny and be held liable more frequently than they currently are. However, she argues that a balance must be struck between a manufacturer's over- and under-exposure to liability in order to preserve both the health of the industry and the well-being of its customers.⁶⁹ Reduced liability of manufacturers may stimulate

⁶⁶ Melinda Florina Lohmann, “Liability Issues Concerning Self-Driving Vehicles,” *European Journal of Risk Regulation* 7, no. 2 (2016): 337.

⁶⁷ UK Department of Transport, 18

⁶⁸ Sabine Gless, Emily Silverman, and Thomas Weigend, “If Robots Cause Harm, Who Is to Blame: Self-Driving Cars and Criminal Liability,” *New Criminal Law Review* 19, no. 3 (2016): 435.

⁶⁹ Lohmann, 339.

a wider proliferation of the technology, in turn allowing it to improve more quickly, but may also place more people at risk while those improvements take place. These tensions reveal the inherent challenge in determining who, in a complex and interconnected enterprise such as the automotive industry, might be held responsible for factors that no human anticipated.

The first deadly self-driving car accident occurred in March 2018, when a self-driving car being tested by the ride-sharing company Uber struck and killed a woman named Elaine Herzberg in Arizona. She had been walking her bicycle across the street in the dark, and according to the preliminary National Transportation Safety Board (NTSB) report, the system detected an object six seconds before the crash, but had difficulty classifying it and predicting its trajectory.⁷⁰ The system needed to gauge whether it should discard the reading as a false positive, such as it would if it detected a blowing newspaper, or whether it should activate the brakes. In this case, its inability to classify the object resulted in it taking no action. This has been reported in some outlets as a “software bug,”⁷¹ but the preliminary NTSB report showed that “all aspects of the self-driving system were operating normally at the time of the crash.”⁷² In reality, the system to distinguish false positives from true positives is extraordinarily complex, and a software bug is not the only reason it might make an error. The problem was more likely an issue of “tuning” the system, or perhaps of insufficiently training it with various stimuli.⁷³ Even if it were a bug, the ultimate responsibility for Herzberg’s death would still not be clear: it

⁷⁰ In an ironic twist, it is possible that the stock vehicle’s built-in emergency braking system would have saved Herzberg’s life, but Uber deactivated it to avoid conflicts with its self-driving system. National Transportation Safety Board, *Preliminary Report: Highway HWY18MH010*, accessed May 29, 2018, <https://www.nts.gov/investigations/AccidentReports/Reports/HWY18MH010-prelim.pdf>.

⁷¹ Timothy B. Lee, “Report: Software Bug Led to Death in Uber’s Self-driving Crash,” *Ars Technica*, May 7, 2018, <https://arstechnica.com/tech-policy/2018/05/report-software-bug-led-to-death-in-ubers-self-driving-crash/>.

⁷² NTSB, 3.

⁷³ Devin Coldewey, “Uber Vehicle Reportedly Saw but Ignored Woman it Struck,” *Techcrunch*, May 7, 2018, <https://techcrunch.com/2018/05/07/uber-vehicle-reportedly-saw-but-ignored-woman-it-struck/>.

could lie with the programmer who made the mistake, the manager in charge of the subsystem, the engineer who integrated the system together, the executive who put the system on the street, or the politician who authorized it to be on public streets. If the final report finds it to be a calibration issue, responsibility might lie with the team responsible for training the system, the manager of *that* subsystem, and so on up the chain. The Uber case is not yet settled, but depending on the root cause determination in the final report, its outcome could establish a precedent for many such incidents as self-driving cars proliferate.

In the wake of Herzberg's death, Uber's CEO said that the company's self-driving cars should be considered "student drivers," and that eventually they would be safer than their human counterparts.⁷⁴ While the comparison is thought-provoking, the personification of the vehicles could be an attempt to deflect responsibility from the corporate entity and onto the car, simultaneously seeking patience and perhaps empathy by conjuring familiar images of teenagers struggling to become comfortable behind the wheel. Whether consumers and regulators will accept such comparisons is yet to be seen.

COMPLEXITY AND THE DIFFUSION OF RESPONSIBILITY

Complexity has long had the effect of diffusing responsibility and, at least on some occasions, reducing or eliminating the ascription of liability. In January of 1986, the Space Shuttle *Challenger* broke apart during launch, killing all occupants: six astronauts and one schoolteacher. An engineer who worked on the faulty solid rocket booster, Robert Ebeling, had warned his superiors that the booster was at risk of catastrophic failure due to the cold weather

⁷⁴ Cara Lombardo, "Uber CEO Says Self-Driving Cars Are 'Student Drivers'," *Wall Street Journal*, April 12, 2018, <https://www.wsj.com/articles/uber-ceo-says-self-driving-cars-are-student-drivers-1523538431>.

expected on the launch day.⁷⁵ Despite Ebeling's persistence, the ill-fated launch proceeded. The official investigation of the incident, known as the Rogers Commission, never assigned blame to any person, office, or organization, instead choosing to dub it "an accident rooted in history."⁷⁶ While the commission did find fault with NASA and its contractors' safety procedures and communication protocols, it stopped short of holding anyone directly responsible.⁷⁷ In civil litigation, Roger Boisjoly, another one of the company's engineers and one of Ebeling's coworkers, sued their company,⁷⁸ but his billion-dollar lawsuit was dismissed by a judge who said that Boisjoly could not demonstrate that he had been victimized by the company; he did not have the proper standing.⁷⁹ Ultimately NASA was required to make numerous technical, procedural, and cultural changes in the aftermath of the accident, but there was no retributive action. The bureaucracy itself, and its enormously complex decision-making process, was to blame.

Another field where complexity intersects with life and death is aircraft operation. The most common factors contributing to aircraft accidents and incidents are human factors, according to a study by Edem Tetteh at Purdue University. The study found that 64 percent of accidents between 1960 and 2000 were attributable to pilot error, and only 22 percent to

⁷⁵ President's Commission on the Space Shuttle Challenger Accident., *Report to the President by the Presidential Commission on the Space Shuttle Challenger Accident* (Washington, DC, 1986), 86.

⁷⁶ Rogers Commission Report, 121.

⁷⁷ Rogers Commission Report, 198-200.

⁷⁸ Philip M. Boffey, "Engineer who Opposed Launching Challenger Sues Thiokol for \$1 Billion," *New York Times*, January 29, 1987, <http://www.nytimes.com/1987/01/29/us/engineer-who-opposed-launching-challenger-sues-thiokol-for-1-billion.html>.

⁷⁹ Joyce E. Cutler, "Federal Judge Dismisses Boisjoly's Lawsuits Against Thiokol," *Deseret News*, September 2, 1988, <https://www.deseretnews.com/article/15938/FEDERAL-JUDGE-DISMISSES-BOISJOLYS-LAWSUITS-AGAINST-THIOKOL.html>

mechanical failure.⁸⁰ A second study, performed by Zohreh Nazeri at George Mason University, agreed that human-related factors were the leading cause, but defined human factors more broadly to include procedures and air traffic control issues in addition to pilot-related factors. In addition, she found that pilot-related factors were exacerbated by combination with any other factor—for example, both pilot-related and weather-related factors would be at play in a crash resulting from a pilot's decision to fly through a storm.⁸¹ This magnification of human error is an indication that the complexity of the human-machine cooperative may be at least partly to blame when such incidents and accidents occur.

When human error (or any other cause) leads to an aircraft accident, an investigation is always undertaken to determine root cause. This practice dates to 1938, and seeks to draw from unique tragedies the vital lessons which might prevent another one from ever happening in the same way.⁸² Ordinarily, corrective actions include procedural or technical adjustments, similar to those undertaken by NASA after the Challenger explosion. In very extreme circumstances, where willful negligence played a role, punishment may result. One such example is a 1994 incident where Air Force Lieutenant Colonel Arthur “Bud” Holland performed increasingly unsafe maneuvers with his B-52H Stratofortress over the course of several years. Ultimately, he crashed a jet while performing an unsafe maneuver, killing himself and three others on board. The wing commander, who had ample opportunity to address Holland's behavior prior to the crash, pled guilty at his court martial for dereliction of duty and received a reprimand and a

⁸⁰ Edem G. Tetteh, "Human Factors Analysis of Commercial Aircraft Accidents in the United States: 1960-2000," *IASE Annual Conference Proceedings* (2006): 3.

⁸¹ Zohreh Nazeri, "Cross-Database Analysis to Identify Relationships Between Aircraft Accidents and Incidents" (PhD diss., George Mason University, 2007), 82-83, ProQuest (AAT 3285768).

⁸² Peter Galison, "An Accident of History," in *Atmospheric Flight in the Twentieth Century*, ed. Peter Galison and Alex Roland (London: Kluwer Academic Publishers, 2000), 3-4.

monetary penalty.⁸³ The wing commander, a colonel who had been selected for promotion, never received his star.⁸⁴ Safety expert Alan Diehl alleged that even those modest punishments were only issued because of the publicity of the event, which was widely and repeatedly televised.⁸⁵

When a single individual or entity is responsible for an undesirable action, retribution is a relatively simple matter. By contrast, when a complex system or bureaucracy seems to be at fault, the result can be a diffusion of responsibility. This can in turn result in no blame assigned, and therefore no justice or closure afforded the victims' families, although corrective actions may well still result. From a utilitarian perspective, this may be entirely acceptable: assuming the corrective action is successful, the problem which led to the incident has been fixed, regardless of blame or punishment. But from the retributivist human perspective with a more values- or duty-based ethic, it may feel incomplete or unjust, or at the very least, unfairly beyond the realm of human control.

Events beyond human control are not isolated to bureaucracies or complex technologies. When nature interferes with human affairs, it is often said that an "Act of God" occurred. Such an event may be defined as "something caused naturally, beyond both man's anticipation and control."⁸⁶ The concept finds use in various legal contexts, such as contractual terms and insurance claims. As science and technology advance, however, the boundaries of an Act of God are eroding on two fronts. On one side, the degree to which weather phenomena are "beyond

⁸³ *Washington Post*, "Colonel Pleads Guilty to Dereliction in Crash," May 20, 1995, https://www.washingtonpost.com/archive/politics/1995/05/20/colonel-pleads-guilty-to-dereliction-in-crash/8179f31b-5353-4dc8-9403-d6ea99005e42/?utm_term=.9ec2bb52fb9c.

⁸⁴ Jim Camden, "Air Force Reprimands, Fines Colonel Officer who was in Charge of Flight Operations, Says B-52 Crash Will Haunt Him for Rest of his Life," *Spokane Spokesman-Review*, May 23, 1995, <http://www.spokesman.com/stories/1995/may/23/air-force-reprimands-fines-colonel-officer-who/>

⁸⁵ Alan E. Diehl, *Silent Knights: Blowing the Whistle on Military Accidents and Their Cover-Ups*, (Washington, DC: Brassey's, 2002), 126.

⁸⁶ Jill M. Fraley, "Re-Examining Acts of God," *Pace Environmental Law Review* 27, no. 3 (2010): 669.

man's control" has become a topic of active debate in a world affected by climate change.⁸⁷ On the other side, humans are devising ever more complex technologies, and may soon produce technologies capable of creating outcomes "beyond both man's anticipation and control." Self-driving cars are one possible example of such technology: the investigation of Elaine Herzberg's death may discover that no human could have anticipated or prevented the system's reaction when it detected her. In other words, the complex interaction of the physical world, sensors, and software defies prediction. For such cases, it may be necessary to establish a new term to stand alongside Act of God to explain those outcomes. The remainder of this chapter proposes such a term: *Act of Code*.

ACTS OF CODE

Common wisdom among computer scientists holds that computers do not make mistakes; they do exactly as they are told. Any undesirable behavior must therefore stem from user, designer, or programmer error. At least, that was the case until recently. The rise of learning and self-programming systems may permit computers to make mistakes in much the same way humans do. Like human mistakes, they may sometimes be difficult—or even impossible—to explain.

At their core, complex, sophisticated systems are still prone to traditional software bugs. A programming rule of thumb is that for every 1,000 lines of code, there will be at least one error.⁸⁸ The U.S.'s most advanced fighter aircraft, the F-35 Lightning II, boasts 8 million lines of

⁸⁷ Fraley, 689

⁸⁸ Chad Perrin, "The Danger of Complexity: More Code, More Bugs," *Tech Republic*, February 1, 2010, <https://www.techrepublic.com/blog/it-security/the-danger-of-complexity-more-code-more-bugs/>.

code; Lockheed Martin touts that this is more than a fourfold increase over the F-22 Raptor.⁸⁹ Conservatively, 8 million lines of code would translate to 8,000 software bugs, and logically we can conclude that it would have four times more than any other fighter aircraft. That is to say nothing of the ground-based support systems the jet needs to fly, which contain another 24 million lines,⁹⁰ and therefore 24,000 bugs or more. While few of these systems connect directly to the internet, they all connect to something, and most of them connect to each other. This certainly increases the software's susceptibility to hackers and malicious code, which is of great concern—but bugs alone can also cause a system to behave unexpectedly. For example, in 2005, Toyota issued a recall for 160,000 Prius vehicles because a software bug caused them to stall without warning.⁹¹ A loss of engine power on the highway is dangerous; in a dogfight at 10,000 feet, it would be catastrophic. Because of human error, no large-scale software-based system, including premier fighter aircraft, can be guaranteed to be free of bugs. The same will be true of autonomous weapons.

Along with unanticipated behaviors caused by human-rooted software errors, the complexity inherent in artificially intelligent systems brings unpredictability on its own. As more and more tasks are leveraged upon neural networks and deep learning algorithms, a unique predicament has begun to reveal itself: it is sometimes impossible to know or understand how the systems arrive at their conclusions. The Massachusetts Institute of Technology (MIT) *Technology Review* relates the story of Mount Sinai Hospital in New York, which employed a deep learning algorithm to analyze medical records. It proved more reliable than human doctors

⁸⁹ F-35 Software Development, last accessed Feb 11, 2018, <https://www.f35.com/about/life-cycle/software>.

⁹⁰ Clay Dillow, "Pentagon Report: The F-35 is Still a Mess," *Fortune*, March 10, 2016, <http://fortune.com/2016/03/10/the-f-35-is-still-a-mess/>.

⁹¹ Simson Garfinkel, "History's Worst Software Bugs," *Wired*, November 8, 2005, <https://www.wired.com/2005/11/historys-worst-software-bugs/>.

at predicting ailments like cancer and, to their great surprise, it seemed able to predict schizophrenia as well—something doctors have never had great success in predicting. As much as the doctors would like to know how the system made its predictions, there is currently no reliable way to interrogate the system to determine its rationale. The reason for this discrepancy is that modern machine learning systems essentially program themselves by processing enormous volumes of data with known solutions, and creating their own algorithms to produce solutions from new data sets.⁹² The complexity of the resultant system does not easily yield to human interpretation. While the Defense Advanced Research Projects Agency (DARPA) does have an ongoing program seeking “Explainable Artificial Intelligence,” one major drawback of the outcomes it produces is that the explanation is always, by necessity, simplified. When neural networks are asked to perform extremely complex analysis beyond human capability, an unavoidable consequence is that there may not be a simple way to explain the results.⁹³

Uncertainty about the specific way a given technology works is not entirely unprecedented. For over a century, medical professionals have used inhaled anesthetic drugs to render patients unconscious for surgical procedures. While these drugs are ubiquitous, no one is entirely certain how they work. There are hypotheses to explain their function, but the drugs were identified and came to be trusted through observation of their effects and effectiveness.⁹⁴ The specific reasons they are effective, while still an area of study, remain a mystery. Despite this fact, their efficacy has not been questioned nor their use halted.

⁹² Will Knight, “The Dark Secret at the Heart of AI,” *MIT Technology Review*, April 11, 2017, <https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/>.

⁹³ Knight.

⁹⁴ George Marshour, Stuart Forman, and Jason Campagna, “Mechanisms of General Anesthesia: From Molecules to Mind,” *Best Practice & Research Clinical Anaesthesiology* 19, no. 3 (2005): 349, <https://www.sciencedirect.com/science/article/pii/S1521689605000054>.

Even humans themselves can take actions without being able to explain why, or make decisions without fully understanding the reasons. Humans develop trust in one another by observing their behaviors and responses to stimuli over a period of time. After an indeterminate amount of time, they become comfortable enough to begin assuming that the new person's actions will comport with their expectations: they trust them. If at some point the new acquaintance takes an action outside of those parameters, that trust may be strained or broken. It may be necessary to treat neural networks and artificial intelligence in a similar manner: observe their actions over a period of time to achieve a requisite level of comfort, and place trust in their continued behavior unless such trust is violated.

When enormously complex algorithms do not perform as expected, it may exceed human capacity to determine what happened, and whether it will ever happen again. If the action was anomalous but not significantly harmful, we may be comfortable continuing to use the algorithm and accepting the risk that the undesired behavior will recur. We may also be able to provide the neural network with additional "training" or data sets which will hopefully reduce the likelihood of a recurrence. This may be considered analogous to military members who commit minor errors. Their leadership can accept their quality of work as-is, or subject them to additional training. If, however, an AI system takes an action which is sufficiently damaging, we may elect to cease using the system entirely in favor of a different method. This would be analogous to military members who commit serious crimes; their military service is likely over. AI systems, however, could not be held responsible for "crimes," and so such consequences may not be possible. After thoroughly investigating why the AI system made its error, if no understandable reason is forthcoming and no humans are found to be at fault, investigators may choose to document the incident as an "Act of Code."

The ability to label an accident an Act of Code is not a substitute for due diligence. It would clearly be unacceptable to unleash an artificially intelligent system without performing rigorous tests and evaluations to determine, with as little doubt as possible, how the system would react to all possible stimuli. The notion of Acts of Code merely acknowledges that there will likely be no feasible—and perhaps no possible—way to evaluate the system under every possible scenario, particularly when it has self-learning attributes which lead to its behavior changing over time, like a human's. The burden of due diligence rests with system designers and testers, as well as with those who approve the AI for active use. Certifying bodies will need to produce minimally acceptable standards for such testing, against which the actions of the aforementioned designers, testers, and approvers can be measured in the event of an incident to aid in the determination of whether they applied due diligence. There will always be humans in the causal chain who must mitigate risk to the greatest extent possible, and accept whatever risk cannot be mitigated if those risks can be deemed acceptable. If those humans have applied due diligence, then the risk is accepted not by the individuals, but by the organizations they represent and, by extension, society.

Approving an artificially intelligent system for use could potentially bear some similarity to approving the use of new medications. The Food and Drug Administration oversees a rigorous development and testing process to minimize the danger of unforeseen problems from new drugs.⁹⁵ Despite the due diligence applied, such measures occasionally fail to detect harmful aspects of the drugs, and those drugs are withdrawn from the market. Such was the case for the weight-loss drug commonly known as “fen-phen” (fenfluramine/phentermine), which was shown

⁹⁵ U.S. Food and Drug Administration, “The Drug Development Process,” last accessed Feb 11, 2018, <https://www.fda.gov/ForPatients/Approvals/Drugs/ucm405622.htm#Approval>.

to increase the risk of heart problems and was removed from the market in 1997.⁹⁶ One major difference between drug approvals and AI approvals stems from the fact that while the chemistry of a given drug does not change, self-learning AI systems will change over time, with the result that the fielded system may diverge substantially from the system as tested, necessitating periodic retesting. The case of fen-phen highlights another potential pitfall: fen-phen was marketed as an easy solution to weight loss, and so an enthusiastic market and profit-hungry doctors greatly increased its reach and, by extension, its harm.⁹⁷ If AI comes to be seen as a “silver bullet” for various problems, such systems could become extremely widespread before researchers and regulators can establish their full impact. This idea is in tension with the suggestion highlighted above, that proliferation of AI-driven systems such as self-driving cars might accelerate their improvement, and so manufacturers’ liability should be limited in order to encourage such proliferation. Ultimately, the systems’ complexity and inscrutability could make it impossible to know which approach will be more successful in any particular case.

CONCLUSION

Unless AI systems truly become self-aware and human society deems them able to stand for their own crimes, a human will always have to accept the risk of their implementation. Because of their extraordinary complexity, no human could be expected to anticipate every outcome of such systems’ interaction with the boundless stimuli of the world. A better approach would be to optimize the system to reduce risk to an acceptable level.⁹⁸ To accomplish this, there

⁹⁶ Gina Kolata, “How Fen-Phen, a Diet ‘Miracle,’ Rose and Fell,” *New York Times*, Sept 23, 1997, <http://www.nytimes.com/1997/09/23/science/how-fen-phen-a-diet-miracle-rose-and-fell.html>.

⁹⁷ Kolata.

⁹⁸ Dave Gershgorin, “The Case Against Understanding Why AI Makes Decisions,” *Quartz*, January 31, 2018, <https://qz.com/1192977/the-case-against-understanding-why-ai-makes-decisions/>.

must be minimum standards and formal processes in place to provide for the certification of such systems, and to protect the humans who accept the risk from liability for truly unforeseeable outcomes. Nowhere is this more relevant than for military technology, where such artificially intelligent systems will not just be charged with *preventing* injury or death, but perhaps with intentionally *causing* injury or death. The margins for error must be as slim as possible. The 1961 exhortation to have lawyers examine computer systems looking for potential problems may come to fruition after all; it will likely be lawyers who determine when the threshold for due diligence has been met, which would push any further unexpected outcomes into the realm of Acts of Code.

Because of the tremendous benefits promised by AI, it would be unreasonable to refrain from its use for the sole reason that its actions cannot always be explained, or that its behavior might change over time. The same could be said of human soldiers with rifles in their hands, and human pilots with massive devastation at their fingertips. As Gless, Silverman, and Weigend said in their study about self-driving cars, “If society embraces the convenience, the opportunities, and the safety assurances associated with self-driving cars, it should also be willing to accept the fact that (possibly very rare cases of) unexpected actions of robots will lead to (generally) foreseeable harm to random victims.”⁹⁹ The sentiments translate well to military technology. While no error is desirable, human error is already a factor and, as discussed in Chapter 1, automation and autonomy may reduce overall errors, and generally provide more advantages than drawbacks. There will come a time when military necessity compels us to develop trust with artificially intelligent systems, and to accept that occasionally an unexpected Act of Code will result.

⁹⁹ Gless et al., 436.

Chapter 3

Distancing and the Humanity of War

INTRODUCTION

“To fight from a distance is instinctive in man. From the first day he has worked to this end, and he continues to do so.”

—Ardant du Picq¹⁰⁰

The first chapter of this work focused on the reasons national decision makers and military leaders might find the notion of autonomous military aviation compelling, perhaps even vital. The second chapter discussed some of the legal and cultural concerns likely to arise in the implementation of such technology, under the reasonable assumption that the technology, even if it surpasses human capability, will never be infallible. This chapter will explore a more philosophical aspect of autonomous military aviation: If the technology exists, and can reliably perform a given task better than a human pilot could, are there reasons that, nevertheless, it *should not* be employed in combat?

The notion of autonomous “killing machines” is not as significant a departure from the current state of affairs as it may seem. As nineteenth century military theorist Ardant du Picq alluded in the quotation above, human society has long sought to place distance between the person doing the killing and the person being killed, whether in malice or in warfare. Hands gave way to rocks, which gave way to swords, then arrows, then firearms, then artillery, then airplanes armed with missiles and bombs. Today, a soldier, sailor, airman, or marine can visit death upon scores of faceless “targets” with the push of a button. Many of the weapon systems they use—be they guided artillery rounds, cruise missiles, or smart bombs—perform active maneuvers in their final moments in a manner which can only be described as “autonomous.” Many such systems,

¹⁰⁰ Dave Grossman, *On Killing* (Boston: Little, Brown and Company, 1995), 107.

once released, have no provision for recall or cancellation of their task. The last human action, as measured by the release of the weapon, precedes the kill itself by seconds, minutes, or even hours.¹⁰¹ A system which is controlled by artificial intelligence (AI) might automate more of the presently human role, creating more physical, temporal, and psychological distance, but would not be a departure from the trend identified by du Picq.

In a world which already has such highly-automated weapons, the additional step to incorporate AI might be considered trivial, but it has already inspired significant opposition. Notably, the Future of Life Institute, a consortium of leading AI researchers, roboticists, and theorists including Elon Musk, Noam Chomsky, and the late Stephen Hawking penned an open letter in 2015 exhorting governments to institute a ban on autonomous weapons—which they defined as those able to make the decision to kill humans based on pre-established parameters, but independently of direct human action. They described autonomous weapons as “the third revolution in warfare, after gunpowder and nuclear arms.”¹⁰² They drew a distinction between such weapons and cruise missiles, but both systems are given instructions prior to launch, and both take independent action to kill without further human input. We might ask whether there is a practical and ethical difference between launching a cruise missile to destroy a target 60 minutes later, and launching an autonomous platform designed to systematically locate, identify, and destroy its target, perhaps days later. Rationally, the only difference is the complexity of the system and the instructions given it—by a human. Nevertheless, the societal response that it

¹⁰¹ The Tomahawk cruise missile (Block II TLAM-A variant) travels at 550 mph for a range of up to 1500 mi. It should be noted that in this case, the weapon can be redirected while in flight if necessary, but no further instructions after launch are required. U.S. Navy, “U.S. Navy Fact File - Tomahawk Cruise Missile,” last modified April 10, 2017, http://www.navy.mil/navydata/fact_display.asp?cid=2200&tid=1300&ct=2.

¹⁰² Stuart Russell et al., “Autonomous Weapons: An Open Letter from AI & Robotics Researchers,” Future of Life Institute, accessed April 13, 2018, <https://futureoflife.org/open-letter-autonomous-weapons/>.

“feels” different, along with the gravity of the scientists’ warning, warrants further exploration. If killing in this manner is somehow different, what makes it so? Does automating the killing further than it has already been automated cross a moral or ethical threshold? Might different ethical analyses arrive at different conclusions? This chapter employs various schools of ethical thought to explore these questions.

The Future of Life Institute theorists’ open letter cautions that the relative ease of killing with such weapons might lower the threshold for violence.¹⁰³ This is one possible result of “distancing” in warfare, but it is not the only one. After a brief primer on ethics and Just War, this chapter will consider four potential outcomes of such distancing: one “light side,” and three “dark sides.” First, it will revisit the “light side” explored in Chapter 1: the tactical advantage such technologies could provide by striking from a distance; they could win battles and save military lives by keeping troops far from harm. For the first of three “dark sides,” it will consider the aforementioned warning from the researchers: the potential that national leaders will more readily resort to violence. Next, it will discuss the potential for an unhealthy relationship between warriors and the violence they unleash. Finally, it will consider a possible “backlash” effect: an increased risk to civilians from terrorism if military combatants become unreachable. The chapter will conclude that, though there are tangible and appealing benefits of creating distance, decision makers must take deliberate steps to preserve the humanity of war, perhaps even to include refraining from the use of available technologies which place that humanity out of reach.¹⁰⁴

¹⁰³ Russell et al.

¹⁰⁴ “Humanity” is used here in a broad sense, without specific reference to international humanitarian law, human rights, or any other formal construct. It refers to war’s human element, and to its significance as a human activity; perhaps, as Gen George S. Patton is said to have observed, its supreme significance: “Compared to war, all other forms of human endeavor shrink to insignificance. God, how I love it!” Forrest C. Pogue, “It’s Hard to Decide Whether Patton Was Drunk From War,” Review of *Patton: Ordeal and*

The concepts in this chapter are more speculative and philosophical than those in previous chapters, and so the conclusions will lean more heavily on prior philosophical thought, including the Just War tradition. As in most explorations of morality, “right” and “wrong” are elusive extremes, and so this chapter makes a supported qualitative judgment about the moral worth of the employment of autonomous military aircraft. It is the reader’s prerogative, and history’s, to conclude differently.

INTRODUCTION TO ETHICS THEORIES AND JUST WAR PRIMER

Ethics studies generally fall into three broad categories: consequentialism, deontology, and virtue ethics. Consequentialism, also known as utilitarianism, holds that the outcome of an action is what determines its goodness—stealing a loaf of bread may be good if done to feed one’s starving family, for example.¹⁰⁵ Deontology, by contrast, is focused on the intent of the action, and specifically whether it was performed out of a sense of duty. Under deontology, even stealing the loaf of bread to feed one’s family might not be good if it were motivated by a desire to be heroic.¹⁰⁶ Virtue ethics, meanwhile, contends that actions or characteristics have inherent value—returning to the bread example, familial devotion might be considered an intrinsically virtuous quality, rendering the action good; or, alternatively, unlawfulness might be considered a vice, rendering the action bad. Virtue ethics depends heavily on the standard applied, and so

Triumph, by Ivan Obolensky, *The Washington Post* (December 18, 1964): A18. ProQuest, <https://search-proquest-com.usnwc.idm.oclc.org/docview/142086667?pq-origsite=summon>.

¹⁰⁵ *Stanford Encyclopedia of Philosophy*, s.v. “Consequentialism,” accessed April 13, 2018, <https://plato.stanford.edu/entries/consequentialism/>.

¹⁰⁶ *Stanford Encyclopedia of Philosophy*, s.v. “Deontological Ethics,” accessed April 13, 2018, <https://plato.stanford.edu/entries/ethics-deontological/>.

universal agreement among various systems is rare.¹⁰⁷ These broad categories of ethical thought are almost never studied in isolation; they are more often used as lenses to examine ethical action or decision-making, and ethicists will often rely on all three to analyze a question.¹⁰⁸

When considering the ethics of warfare, all three schools of ethical thought contribute elements to what has become known as Just War theory. Just War theory is not an official legal standard, nor is it the product of a single philosopher or school of ethics; rather, it is the aggregate of philosophical thought and international norms concerning warfare established over centuries of human conflict. Two of its major components are *jus ad bellum* (the law of going to war) and *jus in bello* (the law of the conduct of war). *Jus ad bellum* itself contains several provisions: war must be waged by a legitimate authority (a head of state); war must be waged for a “just cause,” such as self-defense or to stop an atrocity; it must be waged with “right intention,” meaning that leaders cannot have an ulterior motive such as profit or plunder in mind when waging war; the war must have a reasonable chance of success; the war will cause more good than bad (proportionality);¹⁰⁹ and finally, all other options must have been exhausted prior to waging war (“last resort”).¹¹⁰ *Jus in bello* has several components of its own, which all seek to maximize the safety of noncombatants. They are distinction (or discrimination), which holds that only combatants should be the subject of military attacks; necessity, which states that collateral

¹⁰⁷ *Stanford Encyclopedia of Philosophy*, s.v. "Virtue Ethics," accessed April 13, 2018, <https://plato.stanford.edu/entries/ethics-virtue/>.

¹⁰⁸ "Each of the [ethics] approaches can make room for virtues, consequences, *and* rules. Indeed, any plausible normative ethical theory will have something to say about all three." *Stanford Encyclopedia of Philosophy*, s.v. "Virtue Ethics," accessed April 24, 2018, <https://plato.stanford.edu/entries/ethics-virtue/>

¹⁰⁹ Proportionality is a consideration for both *jus ad bellum* and *jus in bello*; for *jus ad bellum*, it seeks to determine whether going to war is preferable to doing nothing; for *jus in bello*, it seeks to determine the propriety of individual actions.

¹¹⁰ *Stanford Encyclopedia of Philosophy*, s.v. "War: Jus ad Bellum," accessed April 13, 2018, <https://plato.stanford.edu/entries/war/#JusAdBell>.

damage to noncombatants is only permissible if there is no alternative; and proportionality, which mandates that if collateral damage to noncombatants cannot be avoided, the attack should only proceed if the benefit of the attack exceeds that collateral damage.¹¹¹

Because *jus ad bellum* primarily concerns justification for going to war and not its conduct, the weaponry available to fight the war—including autonomous weapons—should not generally affect the determination of whether a war is just. (However, as we will see, military capabilities do affect the feasibility of a proportional victory under *jus ad bellum*, which contributes to a nation's decision to go to war.) The weaponry does, however, have great bearing on *jus in bello*. If an autonomous weapon allows greatly increased precision, it will also permit more effective distinction, a foundational concern of *jus in bello*.¹¹² In that case, minimum force could theoretically be used to accomplish objectives with little or no collateral damage, enhancing a nation's ability to fight within the bounds of *jus in bello*.

THE “LIGHT SIDE” OF DISTANCING: SAFETY FOR FRIENDLY FORCES

Distance creates safety. If an army can accomplish its objectives from outside the range of its enemy's weapons, then its soldiers are theoretically assured of survival. As already noted, the evolution of weapons of war has allowed warriors to be farther and farther from those who would do them harm, while still enabling them to inflict harm. At the forefront of this trend, remotely piloted aircraft allow pilots to actively fight a war thousands of miles away and still sleep in their own beds every night. AI-controlled warplanes would theoretically not even require

¹¹¹ *Stanford Encyclopedia of Philosophy*, s.v. "War: Jus in Bello," accessed April 13, 2018, <https://plato.stanford.edu/entries/war/#JusBell>.

¹¹² Distinction is arguably the foundational concern of *jus in bello*, insofar as protection of non-combatants is *jus in bello*'s purpose.

human attention during their operation. In the extreme, AI could prosecute a war on behalf of the military forces who provided it with parameters, but entirely separate from them, allowing them to incur no physical risk.

Until such technology becomes a reality, automation has great potential to lessen the human cost of warfare while fighting it more effectively. Automated systems may be able to perform certain combat tasks more quickly and effectively than humans, removing those humans from harm and providing an advantage which may eventually become essential to victory. If these capabilities prove practicable, it could be considered irresponsible and dangerous to entrust a life-and-death mission to a human who is more fallible than a machine, or to risk a human life for a task which AI could perform better. In this case, the increased mission effectiveness and the concomitant reduction of risk justify the means used to attain them.

Consequentialist ethics would contend that the mechanics of a military weapon are immaterial to the “rightness” of using that weapon. What matters is the outcome of the weapon’s employment. If we accept this premise, and if we have moral confidence and trust in the military personnel who do violence on our behalf (and the politicians who hold the reins), then equipping them with AI-enabled weapons would not change the moral value of their actions. If an automated weapon can more effectively achieve the just aims of a warring state, its use would be morally defensible from this perspective.

This utilitarian perspective requires an important assumption: AI weaponry will be used toward the same ends as traditional weapons. That is, its mere existence will not have any influence on military or government leaders’ ethical decisions about war, or about specific targets. If such weaponry *does* influence those decisions, then the assertion that a warring nation’s armament does not impact the morality of its actions becomes tenuous. To understand

this further, it is worthwhile to examine some of the factors which affect leaders' decision-making.

Military and government leaders are, of course, human, and will not make every decision based on entirely rational considerations. Even if their decision-making were somehow purely rational, their decisions might be based on different criteria than their constituency expects, or on incomplete or even inaccurate information. Two factors in particular which can cloud the rational basis of decisions are subjectivity and passion.

In the dynamic landscape of international relations, leaders must often rely on information which has been vetted and relayed numerous times, and which may have been interpreted subjectively at multiple levels. In turn, the leaders will apply their own subjective interpretation to the information. Such subjectivity need not be malicious or deceptive in nature; it is merely the individual context in which information is interpreted. Additionally, high-ranking members of various governmental organizations, including military services, spend the entirety of their careers advocating for their organizations. They are accustomed to solving problems with the tools at their disposal, regardless of whether there are tools *outside* their control which might be more appropriate. This phenomenon has become known as Maslow's Hammer, after the influential psychologist Abraham Maslow, who said "it is tempting, if the only tool you have is a hammer, to treat everything as if it were a nail."¹¹³ For example, the Air Force might advocate a long-range strike on a target in denied territory, while Special Operations Forces might propose a raid which could accomplish the same objective with less collateral damage. The Air Force was not being short-sighted; it offered the tools at its disposal, as any organization must. On occasion, organizational leaders may also advocate for their own solutions to increase their organization's

¹¹³ Abraham H. Maslow, *The Psychology of Science; A Reconnaissance* (New York: Harper & Row, 1966), 15.

relevance or funding. Institutional imperatives such as these do not relieve the higher-level decision maker of his or her responsibility to be objective, but they can inspire his or her advisors to oversell their organizations' capabilities (perhaps unintentionally), making objective analysis more challenging.

Further muddying the ability of decision-makers to be fully rational is the fact that it is nearly impossible to exclude emotionality from a topic as cataclysmic as war. Nineteenth-century Prussian General Carl von Clausewitz described the nature of war as a “remarkable trinity” of reason, passion, and chance.¹¹⁴ The three elements of his trinity correspond loosely to the government, the populace, and the military, respectively, but all three elements influence all three bodies. Passion undermines reason when making decisions. The passion of the populace can induce leaders to make a decision to appease that passion (particularly in a democratic society, where a leader who countermands the will of the people does so at the risk of his or her position); passion in the branches of the military can exacerbate the prominence of institutional imperatives in recommended courses of action; and passion among governmental leaders can lead to diminished objectivity in the ultimate decision. Even the technological capabilities of the military, while they may be measured in coldly scientific terms such as range, duration, and blast radius, may have a significant impact on how a nation or a military views its ability to address a problem, and may effectively magnify the influence of passion.

Because subjectivity and passion are inescapable elements of human decision-making in matters of great import, it would be imprudent to accept an assumption that military and government leaders always make perfectly rational, unbiased decisions with regard to the use of

¹¹⁴ Carl von Clausewitz, *On War*, trans. and ed. Michael Howard and Peter Paret. (Princeton: Princeton University Press, 1976), 89.

force, or that technological developments cannot affect those decisions. For example, AI weapons could create an institutional imperative for some branch of the military to demonstrate its investment's effectiveness. Alternatively, by providing a new avenue to exert military force, such weapons could eliminate some of the drawbacks of military action which might otherwise constrain a crisis response borne of passion. Because of these possibilities, a consideration of the threshold for entering hostilities, and how technology might impact that threshold, is worthwhile.

THE FIRST “DARK SIDE” OF DISTANCING: LOWERING THE THRESHOLD FOR WAR

When decision makers are evaluating the financial and human costs of a course of military action, they may consider the justification for the action (*jus ad bellum*) in isolation, but it is most likely that military strength and public support will be part of the calculation as well. If the technological capabilities available to the military will translate to few anticipated casualties, public support (passion) may be assumed to be more reliable, and decision makers may be more inclined to resort to violence. Conversely, if they anticipate more casualties, and therefore a possible deterioration of public support due in part to the so-called “CNN effect,”¹¹⁵ they may be less inclined to take military action.

In recent years, the United States has targeted terrorists outside designated combat zones using remotely piloted aircraft (RPAs)—conducting what are commonly, but imprecisely (because of the active human operators), called “drone strikes.” Because the U.S. had no prior presence in those areas, without the RPAs it would have had to deploy special operations teams

¹¹⁵ Eytan Gilboa, "The CNN Effect: The Search for a Communication Theory of International Relations," *Political Communication* 22, no. 1 (February 2005): 27.

or send a piloted aircraft to conduct the mission—both of which would have placed U.S. lives at risk. With the RPAs, the worst-case scenario for the mission (from the American perspective) would be the destruction of an expensive aircraft with no U.S. lives lost. It is conceivable that the existence of RPAs enabled national decision-makers to take actions they would not have taken without that technology; Kenneth Himes, quoting a nonpartisan think tank report, says that drone warfare “could encourage the pursuit of targets ‘that would be deemed not worth pursuing if ... forces had to be put at risk.’”¹¹⁶ Effectively, the technology creates options for political leaders and may enable a lower threshold for military action. AI weapons may be employed in a similar manner as RPAs, or could contribute to an even lower threshold for violence.

This is problematic. The weapons available to fight a war should not have an impact on the *jus ad bellum* decision about whether it is proper to go to war—particularly not in a nation such as the United States, where sufficiency of military power is rarely in question. Similarly, if we accept the consequentialist premise (the “light side,” above) that only the results of an action matter, then the morality of actions taken in a war does not change depending on the weapon. Meanwhile, *jus in bello* does not distinguish the different types of weapons, only the effect they cause.¹¹⁷ If a nation’s decision to go to war hinged on the ability of its weaponry to make that war more palatable, it is possible that considerations other than *jus ad bellum* were incorporated in the decision to fight. In that case, the decision to go to war ought to be rigorously examined.

It may be helpful for decision-makers to remember that the physical and psychological detachment afforded them by their positions does not change the outcome of their decisions. No

¹¹⁶ Stimson Center, “Recommendations and Report on the Task Force on Drone Policy,” co-chairs, John Abizaid and Rosa Brooks; project director, Rachel Stohl (Washington, DC: Stimson Center, 2014), 11, quoted in Kenneth R. Himes, *Drones and the Ethics of Targeted Killing* (Lanham: Rowman & Littlefield, 2016), 151.

¹¹⁷ Certain classes of weapons are prohibited by international law, such as biological weapons. These are prohibited under the Law of Armed Conflict provision against unnecessary suffering.

matter the distance—the range of a rifle, a bomb, or even the psychological distance afforded by the sterility of command—there is no change in the moral value of deciding to kill. As Vietnam veteran and author Karl Marlantes vividly states, “Killing people with Marines is ethically no different from killing people with hatchets.”¹¹⁸ Marlantes goes on to emphasize that politicians must acknowledge their relationship to killing when they decide to go to war: “Only if [leaders] see that *they* are actually doing the killing can they make a more conscious decision... Distancing from the action should not preclude the leaders’ use of their imagination so that they can get into the correct relationship with the decision to wage war. Without this leap of imagination, modern political leaders will not be prepared to think and behave like ethical warriors.”¹¹⁹

U.S. RPA strikes in countries such as Yemen have been justified as self-defense—a permissible reason for violence under *jus ad bellum* if the threat is imminent. While long-standing precedent in customary international law has held that “imminent” means the need for self-defense is “instant, overwhelming and leaving no choice of means and no moment for deliberation,”¹²⁰ the U.S. has more recently promulgated a broader interpretation of imminence. President Barack Obama justified the strikes in Yemen as countering “a continuing and imminent threat” posed by terrorists against the U.S.¹²¹ Similarly, Harold Koh, former chief legal advisor to the U.S. Department of State, coined the term “elongated imminence” to describe the threat posed by terrorists who always seek to do harm to the U.S., although they might not be

¹¹⁸ Karl Marlantes, *What it is Like to Go to War* (New York: Atlantic Monthly Press, 2011), 227.

¹¹⁹ Marlantes, 227.

¹²⁰ Rick Gladstone, “If U.S. Attacks North Korea First, Is That Self-Defense?,” *The New York Times*, August 10, 2017, <https://www.nytimes.com/2017/08/10/world/asia/us-north-korea-preemptive-attack-questions-answers.html>.

¹²¹ Obama, Barack, U.S. President. Address, National Defense University, Fort McNair, Washington, DC, May 23, 2013, <https://obamawhitehouse.archives.gov/the-press-office/2013/05/23/remarks-president-national-defense-university>.

actively engaged in terrorist activities.¹²² Himes synopsized this outlook by saying that because “al-Qaida is always in the midst of plotting terrorist attacks ... the assumption must be made that any member of al-Qaida or associated forces is an imminent threat.”¹²³ He goes on to say of this and similar policies: “[The U.S.] ought to establish norms that we would wish others, including our rivals, to follow. ... Do we wish to live in a world where drones can attack anywhere[?]”¹²⁴ This echoes what 18th century German philosopher Immanuel Kant called the “categorical imperative,” his universally applicable standard for determining the moral worth of an action. The imperative is to “act only according to that maxim whereby you can, at the same time, will that it should become a universal law.”¹²⁵ In this case, that could mean that the U.S., by targeting terrorists in non-warring nations, is implicitly authorizing other nations to do the same within the U.S.’s borders for anyone they consider to be a terrorist.¹²⁶ That would not be a popular notion for U.S. government officials, nor would it be feasible for most nations to undertake such a strike in U.S. skies. However, the categorical imperative would hold that if the U.S. takes such actions, it approves them universally.

Applying Kant’s categorical imperative to future technologies might mean that if the U.S. uses AI to locate and engage adversaries, it licenses other nations, including adversaries, to do the same. Because the technology for AI weapons could conceivably become very accessible to

¹²² Kenneth R. Himes, *Drones and the Ethics of Targeted Killing* (Lanham: Rowman & Littlefield, 2016), 95.

¹²³ Himes, 129.

¹²⁴ Himes, 17.

¹²⁵ Immanuel Kant, *Foundations of the Metaphysics of Morals and Critical Essays*, ed. Robert Paul Wolff, trans. Lewis White Beck (London: MacMillan Pub Co, 1990), 44.

¹²⁶ In order to conduct an attack against a target in a country where the U.S. is not engaged in open hostilities, the U.S. must demonstrate that the country in question is unwilling or unable to prosecute the target itself; while the application of this standard would seemingly reduce the incentive for other countries to conduct comparable operations on U.S. soil (the U.S. is certainly *able* to prosecute such targets within its borders), differing ideologies or viewpoints might render the U.S. *unwilling* to do so, and the countries may feel justified in attempting such an operation.

almost any nation or armed group, the norms established by leading technological and military powers will be important for their future use. Even if the U.S. declines to develop autonomous weapons, it will still have a role in establishing international norms if and when it responds to another nation which does develop and deploy the technology. If the other nation's technology is sufficiently advanced, U.S. response options may be limited, effectively ceding the initiative of establishing norms to an adversary.

Kant argued that the motivation for an action holds its moral worth, a concept at the core of deontology: an action is good if performed out of a sense of duty. In *jus ad bellum*, this aligns with the requirement of "right intent." Nations must enter into warfare, Kant might say, not for national interest, but out of a sense of duty to humanity. If nations truly adhere to the requirements of *jus ad bellum*, the threshold for taking action should not move as technologies evolve. There may, however, be situations which meet the standard of *jus ad bellum*, but were formerly not prosecutable because of a shortfall of technology. The difference between these two scenarios might be very slight. For example, if a nation (Country A) observes that another nation (Country B) is suffering deteriorating conditions under an authoritarian regime, it may evaluate whether it is justified in taking action. If it determines, through careful introspection, that its primary motive for taking action is to stabilize oil prices, then it has met the criteria of legitimate authority and just cause, but not of right intention. It would not, therefore, be a just war. In this scenario, *jus ad bellum* should prevent the nation from taking action. In a similar scenario except with right intention, Country A may meet all criteria of *jus ad bellum*, but find that its military is lacking the capability to sufficiently discriminate between combatants and civilians; this lack of distinction would lead to enough collateral damage that it could not be considered proportional.

In this scenario, outwardly very similar to the first, *jus in bello* provides the criteria by which the nation should choose not to act.¹²⁷ The only difference is Country A's motivation.

Now consider that Country A acquires a technology which will allow it to effectively discriminate between combatants and noncombatants, with little or no collateral damage. Under the first scenario (concern for oil prices), an attack would still be impermissible by *jus ad bellum*, but under the second, hostilities would now be considered just. Technological development can, therefore, provide a formerly lacking means to a just end; the threshold for action may have changed, but that does not mean that the bar for justified violence has been lowered. It remains possible, of course, for nations like Country A to embark on unjust campaigns with the benefit of new technology without meeting all the criteria of *jus ad bellum*. Because of the subtleties and subjectivities involved, the difference between using technology to wage an unjust war and using it to render a just war possible may be only distinguishable by historians, lawyers, and ethicists with the benefit of hindsight—and, of course, by the warring leaders themselves. The fact that national decision-makers' true motives are private does not relieve them of the immense responsibility to scrutinize those motives, and to adhere to Just War principles. National leaders are entrusted to make the morally correct life-and-death decisions on their citizens' behalf, and on behalf of humanity itself.

¹²⁷ This could also be considered proportionality under *jus ad bellum*; the nation considering action could not guarantee that the good outcomes of action would sufficiently outweigh the adverse ones. *Jus ad bellum*'s version of proportionality implicitly accounts for the feasibility of the action in light of available military capabilities.

THE SECOND “DARK SIDE” OF DISTANCING: AN UNHEALTHY RELATIONSHIP WITH VIOLENCE

If the distance provided by modern weapons can shift politicians’ outlook toward violence, it must have an even greater effect on the warriors themselves, where technology serves as their interface to death. In some cases, physical and psychological distance allow the warrior to be nearly unaffected by the violence. In other cases, the very technologies which allow physical distance and consequent safety can cause a drastic collapse of psychological distance, with damaging effects. Determining the “appropriate” psychological distance is a challenging task, as is establishing what a healthy relationship with violence should be, if it exists.

Lieutenant Colonel Dave Grossman, as part of his volume *On Killing*, performs an extensive analysis of the psychology of killing at a distance, beginning with the great distances held by those who drop bombs and launch cruise missiles, and progressing incrementally to the grisly intimacy of killing with bare hands. He observes that the closer the distance, the more reluctant humans are to kill, and conversely, adding distance makes it “easier” for humans to kill one another. He also concluded that greater distance results in measurably fewer psychological problems for the aggressor. The crews of the planes which dropped atomic bombs on Japan, where there were hundreds of thousands of casualties, experienced no documented psychological problems, he said. In fact, he notes, “In years of research ... I [have not] found a single instance of psychiatric trauma associated with [long-distance] killing.”¹²⁸ If true, this may be an example of excessive psychological distance where the purveyors of violence have become unable to empathize.

¹²⁸ Grossman, 108.

A lack of psychological trauma for a nation's warriors may be an appealing feature of distancing weapons, and it could be added to the consequentialist benefits of AI weaponry. But the fact that warriors are avoiding psychological problems (if this is indeed true over the long term) is not the only consideration. It may be that those who are entrusted to do violence on behalf of their nation *should* have a natural aversion to that violence; that aversion limits violence in the world. A lack of overt psychological trauma is also not an indication that one's relationship with violence is "healthy." In fact, in the realm of killing, a lack of trauma could very well be a sign that all is not well; perhaps killing *should* be a traumatic experience. Clausewitz contended that the unpleasantness of war is one of the only things restraining it from maximum violence: "If wars between civilized nations are far less cruel and destructive than wars between savages, the reason lies in the social conditions of the states themselves and in their relationships to one another." Those conditions, he continued, are not part of the nature of war, but only "circumscribe and moderate" war.¹²⁹ If it were not for the individual and collective trauma associated with violence, states might have a reduced incentive to limit war's cruelty and destruction.

In general, technological advances have served to move combatants farther apart and create physical and psychological distance. There has been one notable exception to this trend: RPAs. So-called "drones" have continued to add physical distance between warfighters and their targets, but have actually compressed the psychological distance. Because crews may be assigned to observe targets for days or weeks, they can become intimately familiar with their targets' lives and feel very close to them—something which RPA pilot Dave Blair and professional counselor Karen House have termed "Cognitive Combat Intimacy" (CCI). When

¹²⁹ Clausewitz, 76.

they eventually kill the target, high-resolution sensors allow them to witness the aftermath, sometimes including the anguish of the target's family. Psychologists are only beginning to understand the impact on the crews who engage in these types of operations over an unconstrained period of time without opportunity to disengage and decompress.¹³⁰

If RPA operations disrupted the trend of decreasing CCI, artificially intelligent weapons could return to the trend by automating much of what human RPA operators do. In so doing, AI weapons could also eliminate some of the problems which have arisen for RPA crews. Whether they would spawn new problems, however, is an open question. In the extreme, there is a danger that the warfighters who dispatch the weapons might achieve so much distance from the violence they wreak that they will cease to identify with it at all; it may become out of reach for even the most deliberate empathy, reducing the motivation to limit violence described by Clausewitz. In the case of a cruise missile or a bomb, the person who launched or released it can easily, with minimal introspection, understand the consequences of their actions. By contrast, if an AI-controlled weapon is dispatched to (for example) kill anyone it finds wearing an enemy uniform, the operator may not know whether anyone was killed, let alone how many. The operator may not even be aware of what the platform is programmed to do. Conversely, the person who programs the platform might not be involved at all with its release at a much later date. There might be no humans left who feel that they were directly responsible for ending another human life.

It is also possible, moreover, that operators may experience problems which are difficult to foresee, similar to the experience of the RPA community. For example, those who monitor AI-

¹³⁰ Dave Blair and Karen House, "Avengers with Wrath: Moral Agency and Trauma Prevention for Remote Warriors," Lawfare, November 12, 2017, 1, <https://www.lawfareblog.com/avengers-wrath-moral-agency-and-trauma-prevention-remote-warriors>.

controlled weapons might feel that they are accountable for violence outside their control; as discussed in Chapter 2, it is even possible that the weapon might take actions which neither the operator nor anyone else can fully explain. The operator of a weapon of this sort may feel like a “responsible spectator” to untethered and literally inhuman violence. That lack of control in the face of real violence might create a whole new spate of unpredictable psychological problems.

To counter what he perceives as a problematic lack of empathy resulting from cultural, operational, and technological factors, Karl Marlantes advocates a deliberate emphasis on empathy at all levels. He encourages warfighters to make a deliberate effort to understand and deal with the human cost of their actions as quickly as possible after the action itself, so as to avoid long-term psychological problems when the realization eventually surfaces, and to avoid a troubling “blurring” of the line between war and peace.¹³¹ The emphasis on empathy must continue at all echelons, perhaps most importantly at the level of national decision makers.¹³² He maintains that the human understanding of violence is vital. “We must come to grips with consciously trying to set straight [the] imbalance of modern warfare. What is at stake is not only the psyche of each young fighter but our humanity.”¹³³

This outlook—pursuing empathy in warfare for the sake of humanity itself—can be seen as an expression of virtue ethics, which holds that there are inherently good and bad qualities of character, which in turn lead a person to take definably good and bad actions. Aristotle, in Book 2 of his *Nicomachean Ethics*, describes virtuous qualities as the ideal position along a spectrum between two vices. For example, he says that “courage” is the virtue, or ideal, which falls

¹³¹ Marlantes, 25.

¹³² Marlantes, 227.

¹³³ Marlantes, 19.

between the two extreme vices of “rashness” and “cowardice.”¹³⁴ To extend Aristotle’s model, “empathy” might be seen as a virtuous ideal between the extremes of “heartlessness” and “overidentification.” In highly automated warfare, empathy may slip out of reach. Must there be a certain amount of direct human control in the prosecution of violence—for both national leaders and soldiers—to maintain a virtuous ideal of empathy? It might be argued that individual virtues are not important in a war; after all, military members subordinate their will to that of the state, so perhaps only the virtues of the state matter.¹³⁵ But the virtues of a state are nothing but the aggregate virtues of its people, and especially the virtues and actions of those who do violence on the state’s behalf. In order to preserve a balance between the virtues of the individual and those of the state, Marlantes contends that “individuality must not be suppressed even though individual action is subordinated.”¹³⁶ That individuality, and the virtuous empathy it must embody, can result in psychological hazard for the individual. But without it, empathy—and with it, the humanity of warfare—could be lost. It may be extremely difficult for the complex system of engineers, authorizing officials, operators, and leaders to deliberately preserve enough empathy to use AI weapons in a manner which preserves the humanity of the warring state.

THE THIRD “DARK SIDE” OF DISTANCING: TERRORISM

When the United States confronted Saddam Hussein’s invasion of Kuwait in 1991, it did so with overwhelming power, and demonstrated to the world that American military might stood alone at the pinnacle of the post-Cold War world order. In the minds of many U.S. leaders, it

¹³⁴ Aristotle, *Ethics* (Penguin Books: London, England, 1976), 104.

¹³⁵ To further complicate matters, in the realm of AI-enhanced warfare, both the individual and the state will also have subjugated their will to those of algorithms.

¹³⁶ Marlantes, 141.

dispelled the ghost of the Vietnam quagmire. But, in a stirring reminder of Clausewitz's dictum that "in war the result is never final,"¹³⁷ the triumph in the Gulf War was an early step down the road to the terrorist attacks of September 11, 2001, and to U.S. embroilment in counterterrorism campaigns around the world. By that measure, the "end" of the Gulf War has yet to be seen, and one result of the U.S. military's overwhelming power was the inspiration for an attack which demonstrated *increased* risk to the citizens the military protects.¹³⁸

The ability to prosecute a war with seeming invulnerability can inspire resentment among the people on whom the weapons are fired. Himes notes that "drones, because they strike with impunity, suddenly, and almost anywhere, appear to have a singular ability to terrorize not only legitimate targets but people in general within a region under surveillance. Drone strikes also anger those concerned with border sovereignty. ... As a result there has been a significant rise in antagonism toward the United States."¹³⁹ That rise in antagonism can easily be warped into recruitment for terrorist organizations, and motivation for further attacks.¹⁴⁰

The counterterrorism campaign itself may, unfortunately, be self-perpetuating. As Grossman points out, not in relation to counterterrorism, but to violence in general: "Each individual who is injured or killed by criminal violence becomes a focal point for further violence on the part of their friends and family."¹⁴¹ We have seen parallel concerns from senior

¹³⁷ Clausewitz, 80.

¹³⁸ Thomas H. Kean, Lee Hamilton, and National Commission on Terrorist Attacks upon the United States, *The 9/11 Commission Report: Final Report of the National Commission on Terrorist Attacks upon the United States* (Washington, DC: National Commission of Terrorist Attacks upon the United States, 2004), 57-59.

¹³⁹ Himes, 18.

¹⁴⁰ At least one commentator alleges that this effect has been overstated: Aqil Shah, "Drone Blowback in Pakistan is a Myth. Here's Why," Washington Post, May 17, 2016, <https://www.washingtonpost.com/news/monkey-cage/wp/2016/05/17/drone-blow-back-in-pakistan-is-a-myth-heres-why/>.

¹⁴¹ Grossman, 330.

politicians that killing terrorists only inspires their friends and family members to take up the cause in revenge, and the technology used may be an exacerbating factor. Nabeel Khoury, a former State Department official in Yemen, in 2013 estimated that “the U.S. generates roughly 40 to 60 new enemies for every AQAP [al Qaeda in the Arabian Peninsula] operative killed by drones.”¹⁴² When the U.S. military is not vulnerable to attack in the war zone because of its tactics and technologies, enemies may seek out other ways to do damage, as the self-professed Islamic State did in 2016 when it published what appeared to be a meticulously assembled “hit list” of U.S. drone pilots’ personal information, and exhorted its followers to “Kill them wherever they are, knock on their doors and behead them, stab them, shoot them in the face or bomb them.”¹⁴³ Such attacks, if perpetrated, would certainly place the families of such pilots, and anyone else in their vicinity, at great risk.

Asymmetry is not an unusual feature of modern conflict. Army historian Conrad Crane said, “There are two approaches to waging war, asymmetric and stupid.”¹⁴⁴ Commanders have always sought an asymmetric advantage, and when they could not achieve it in sheer power, they attempted to gain the upper hand by attacking their enemies’ weaknesses—achieving localized asymmetric warfare. The disadvantaged combatant will seek an advantage in any way it can, and

¹⁴² Khoury was arguing that the U.S. focus on combating terrorism undermines its ability to support the implementation of stable governance: “The global war on terror sill [sic] trumps the prioritization needed for assisting the democratic transition underway. Drone strikes take out a few bad guys to be sure, but they also kill a large number of innocent civilians. Given Yemen’s tribal structure, the U.S. generates roughly forty to sixty new enemies for every AQAP operative killed by drones.”

Nabeel Khoury, “In Yemen, Drones Aren’t a Policy,” *Cairo Review of Global Affairs*, October 23, 2013, <https://www.thecaireview.com/tahrir-forum/in-yemen-drones-arent-a-policy/>.

¹⁴³ Dipesh Gadher and Toby Harnden, “Islamic State Hackers Publish Hit List of US Drone Pilots,” *The Weekend Australian*, May 2, 2016, <https://www.theaustralian.com.au/news/world/islamic-state-hackers-publish-hit-list-of-us-drone-pilots/news-story/3b64879379c008be78f07000deacc4f1>.

¹⁴⁴ Conrad C. Crane, “The Lure of the Strike,” *Parameters: The U.S. Army War College Quarterly* 43, no. 2 (2013), http://www.strategicstudiesinstitute.army.mil/pubs/parameters/issues/Summer_2013/1_Crane_SpecialCommentary.pdf

that is what the U.S.'s enemies continue to do, by identifying and exploiting U.S. vulnerabilities. In his 1996 book *The Clash of Civilizations*, Samuel Huntington said that the then-“quasi-war” between Islamic militants and the West “has largely been a war of terrorism versus air power.”¹⁴⁵ A few years later, Osama bin Laden determined that the U.S., despite its overwhelming military strength, had a key vulnerability which he had the capacity to exploit: the civilian populace. He was able to exploit that weakness on September 11, 2001, to make good on his 1998 fatwa calling for “the murder of any American, anywhere on earth.”¹⁴⁶

Artificially intelligent weapons may not significantly alter the dynamic of the U.S. asymmetric advantage. On the receiving end, a missile fired by a remotely piloted aircraft might be indistinguishable from one fired by an AI-piloted aircraft. But the seemingly impersonal, impervious nature of the attack will likely continue to inspire survivors to seek revenge in their own asymmetric ways.

CONCLUSION

Good reasons exist to automate as many functions of warfare as possible. The ability to bring greater capability with less risk to the force makes ever-more-advanced weapons appealing for national and military leaders. However, before adopting such weapons, those leaders should ensure that they have evaluated more than just the utilitarian value of the capability. One shortfall of consequentialism is the difficulty of predicting the *long-term* consequences of an action. In the short term, an action may cause some morally good outcome. But in the long term, complexity and a consistent human failure to accurately predict the future make the proposition

¹⁴⁵ Samuel P. Huntington, *The Clash of Civilizations and the Remaking of World Order* (New York: Touchstone, 1997), 217.

¹⁴⁶ *9/11 Commission Report*, 47.

much less secure. This greatly complicates decision-makers' ability to consider the Just War principle of proportionality; with imperfect knowledge of the future, the outcomes of military action cannot easily be weighed against inaction.

Military leaders must examine whether their willingness—perhaps even eagerness—to use a capability will affect their decision making when deciding whether to use military force. They will need to deliberately separate *jus ad bellum*—the principles on which a war will be waged—from *jus in bello*—consideration of the employment of weaponry. Only after determining that their action is just—including being guided by an intention which has not been influenced by available technology—should they consider military options.

They should also be mindful of the potential impact on the men and women who will wield the weapons. Will they become so detached from the violence their weapons create that the horrors of war cease to be a deterrent to waging it? Will their detachment “trickle up” the echelons of command, and infiltrate military culture over the next generation, changing the way our military, our society, and our leadership view war? Or will they find the burden of being a responsible spectator to automated violence too much to bear, wearing down the human element of the nation's military?

Finally, leaders should consider the way these weapons will be perceived by the populations on whom they are used. If a new class of weapons is viewed as unmartial, uncivilized, or inhumane, the use of that class of weapons risks inflaming public passion against their bearers—potentially prolonging the conflict, lowering the chances of stable post-war operations, or even inspiring terrorism. In that case, leaders must be open to refraining from the use of weapons likely to produce undesirable long-term effects, or reconsider the nation's position in the conflict.

In summary, leaders must strive to preserve the humanity of war. When armies charged against one another with swords, there was no escaping the human element. Even bomber pilots and submariners, as protected from the enemy as they are, undertake significant risk to their own lives to perform their missions—a distinctly humanizing experience. As technology advances, that traditional dynamic shifts. RPA pilots work from the safety of their home bases, and we are seeing that instead of putting their lives on the line, they may be putting their mental health on the line—a vulnerable position. Even that vulnerability may vanish in an era of AI weapons, when operators may only establish some criteria for their weapons and release them to perform their tasks, allowing the operators to be both physically and psychologically insulated. As has never been necessary before, warriors and politicians alike must consciously remind themselves of the terrible cost of war, and moderate their willingness to approach it accordingly.

Conclusion

Because of automation, one bomb can destroy what used to take 9,000 bombs.¹⁴⁷ Because of automation, pilots can fly and land astonishingly complex, high-performance aircraft in relative safety—and the planes will even save the pilots' lives from time to time. Because of automation, nations today can come closer than ever to achieving the Just War ideals of distinction and proportionality. And, in the near future, automation may be the only thing which can prevent an enemy from harming the American people. With such benefits, incorporating automation into military technology would seem to be an absolute practical, moral, and ethical necessity. But consequentialism has its limits. One such limit is the persistent human inability to predict long-term consequences.

Technology historian Melvin Kranzberg contended that “technology is neither good nor bad; nor is it neutral.”¹⁴⁸ By this he meant that no technology can be evaluated out of the context of its use, and additionally that as time passes, a technology's impact on the world may change in unexpected and unforeseen ways. One example he provided was the pesticide DDT, which is vilified in the U.S. because of its adverse environmental effects, but exalted in India where it was instrumental in dramatically controlling malaria.¹⁴⁹ Technology cannot be neutral, Kranzberg maintains, because its impact on the world is far from neutral, but the moral worth of technology comes from its intersection with the human dimension of its application. AI is an extraordinarily promising technology for thousands of functions, but the unknowable long-term implications of its military use render a value judgment based on outcomes alone impossible.

¹⁴⁷ Michael R. Gordon and Bernard E. Trainor, *The Generals' War: The Inside Story of the Conflict in the Gulf* (Boston: Little, Brown, 1995), 189.

¹⁴⁸ Melvin Kranzberg, “Kranzberg's Laws,” *Technology and Culture* 27, no. 3 (1986), 545.

¹⁴⁹ Kranzberg, 546.

Inasmuch as consequentialism is based on projections of the future, it fosters a bias toward unrealistically optimistic prediction and insufficient consideration of shortcomings and their repercussions. In reality, some of the projected advantages of automation may never come to fruition—or they may be rendered moot by drawbacks. Meg Leta Jones, assistant professor of Communication, Culture, and Technology at Georgetown University, contends that highly effective automation can cause operators' skills and vigilance to atrophy, foster more trust in the system than is warranted, and leave the operator with only the most difficult tasks—plus the added task of supervising the automated system and recovering from any errors it makes. Thus she asserts a major irony of automation: “the more advanced and reliable the automation, the more important the human operator must be.”¹⁵⁰ That seemingly inverted relationship does not bode well for a system expected to operate without human intervention. In that case, the operator's skills and vigilance become irrelevant, trust in the system must become absolute, and there is no human ability to help the system recover from errors.

Another problem of the strictly consequentialist approach is that it suggests that only ends have moral value, not means. When only the ends establish the morality of an action, the means may not be properly scrutinized. The trend over the history of warfare has been to increase one's own distance from the enemy using technology in order to increase physical safety while still achieving the same ends—the destruction of the enemy. A simple consequentialist approach would find no fault with this arrangement. But if that distance created by that technology is corrupting in some way, or damaging to the operators, or dangerous to the citizenry, or if the simple availability of the new means can insidiously change the government's

¹⁵⁰ Meg Leta Jones, "The Ironies of Automation Law: Tying Policy Knots with Fair Automation Practices Principles," *Vanderbilt Journal of Entertainment and Technology Law* 18, no. 1 (2015): 91.

calculus about the use of force, then the simple equation of ends justifying means must be called into question. One way that technologically-achieved distance could be a corrupting influence is through the cultivation of a transfer of moral agency.

When machines go to war without direct human oversight, it may seem as though they assume moral agency for their actions. The machines cannot accept that agency, of course, but the moral responsibility may still become diffused—to the entire chain of command, or even to a diaspora of engineers, programmers, purchasers, authorizers, and elected officials. No matter how diffuse it seems, the responsibility cannot and must not be abdicated entirely. Because automated systems are created by humans, they necessarily inherit some of our flawed nature. The Congressional report on the *Challenger* explosion asserted that “to dissect and interrogate relentlessly projects and programs that bring home repeated A’s for achievement and accomplishment... is a reflection of respect for the human fallibility that we all possess.”¹⁵¹ A failure to continue to rigorously “dissect and interrogate” autonomous systems to verify that they are still trustworthy would be an abdication of responsibility by those who approve and employ those systems. This will be especially important if (or when) the systems seem to need no oversight at all.

Highly competent machines may do unexpected and disastrous things—and we may term those unexpected things Acts of Code in an effort to absolve any humans of responsibility. But even if all humans involved applied due diligence, some circumstances will demand that a human take responsibility. If an autonomous weapon commits a war crime, a human must answer for that crime—just as a commanding officer must often accept responsibility for the

¹⁵¹ U.S. House of Representatives, *Investigation of the Challenger Accident*, Committee on Science and Technology, Washington, DC: October 29, 1986, 7.

serious transgressions of those under his or her command. To accept Acts of Code as wholly absolving humans of responsibility would be to ignore that humans are the source of and reason for all the violence of war. Norbert Wiener, the father of cybernetics, declared that “to throw the problem of [our] responsibility on the machine, whether it can learn or not, is to cast [our] responsibility to the winds, and to find it coming back seated on the whirlwind.”¹⁵² The notion of an Act of Code may be used as part of an explanation of what happened, but not to relieve the human element in a system of the burden of accountability.

On the very first page of *On War*, Clausewitz defines war as “an act of force to compel our enemy to do our will.”¹⁵³ Notably absent from his description is the destruction of enemy weaponry, or the deployment of superior firepower; his fundamental definition of war centers on the human element. Proponents of automated warfare would do well to remember that all conflict begins and ends in the human mind—no matter how automated or roboticized the front line becomes, it is humans who care about the outcome, not machines. Naval theorist Wayne P. Hughes reminds his readers that in all naval operations, “The seat of purpose is on the land,”¹⁵⁴ meaning that naval operations, while they are vital, ultimately support a land-based objective. In an age of autonomous weapons, Hughes’s dictum may be adapted to produce an AI corollary: “the seat of purpose is flesh and blood.”

The capabilities and the concerns highlighted in this paper are by no means limited to “Airpower.” Autonomy may soon have the capability of replacing not just pilots, but sailors, submariners, tank drivers, satellite drivers, and cyber operators as well—and perhaps someday

¹⁵² Norbert Wiener, *The Human Use of Human Beings: Cybernetics and Society* (New York: Avon Books, 1954), 185, quoted in Jones, 113.

¹⁵³ Clausewitz, 75.

¹⁵⁴ Wayne P. Hughes, *Fleet Tactics and Coastal Combat* (Annapolis: Naval Institute Press, 2000), 34.

even the rifleman. Whichever nation finds itself on the leading edge of technological change will assume the obligation of establishing precedent. Just as the Soviets established that orbital overflight does not constitute a violation of sovereign airspace when they launched Sputnik in 1957, the first nation to use autonomous weapons on the battlefield may establish how those weapons are used and viewed by the world for the foreseeable future.¹⁵⁵ If the U.S. wishes to have a hand in establishing norms for autonomous weaponry—limiting it to solely defensive applications, for instance—it must invest and integrate accordingly.

War is human conflict. It may be fought with rocks, bombs, or drone swarms. The hand that throws the rock is morally and ethically accountable for what the rock strikes. Autonomous weapons add many more degrees of separation, and exponentially more physical and psychological distance, but they do not change the fundamental relationship between the appalling violence of war and the humans who wage it. Humans invented war, and no technology will relieve them of their ultimate responsibility for it.

¹⁵⁵ Roger D. Launius, "Sputnik 'Was Force for World Peace'," *BBC News*, October 4, 2007, <http://news.bbc.co.uk/2/hi/science/nature/6965717.stm>.

Bibliography

Adams, Eric. "New Navy Tech Makes it Easy to Land on a Carrier. Yes, Easy." *Wired*, August 2, 2016. <https://www.wired.com/2016/08/new-navy-tech-makes-landing-aircraft-carrier-breeze/>.

Aristotle. *Ethics*. Penguin Books: London, England, 1976.

Association of Financial Professionals. *2013 AFP Electronic Payments Survey*. Bethesda, MD, 2013. Accessed May 1, 2018. <https://www.afponline.org/docs/default-source/default-document-library/pub/2013-afp-electronic-payments-report-preview.pdf>.

Astor, Gerald. *The Mighty Eighth*. Dutton Adult, 1997. Quoted in O'Mara, Raymond P. "The Socio-Technical Construction of Precision Bombing: A Study of Shared Control and Cognition by Humans, Machines, and Doctrine During World War II." PhD diss., Massachusetts Institute of Technology, 2011. <http://hdl.handle.net/1721.1/67754>.

AT&T Archives. "Her Right Place." AT&T Tech Channel. Release date July 1, 2013. <http://techchannel.att.com/play-video.cfm/2013/7/1/AT&T-Archives-Her-Right-Place>.

Benusson-Butt, David. "Night Photographs in June–July 1941: A Statistical Analysis." 1941. <https://etherwave.files.wordpress.com/2014/01/butt-report-transcription-tna-pro-air-14-12182.pdf>.

Blackhurst, Rob. "The Air Force Men Who Fly Drones in Afghanistan by Remote Control." *The Telegraph*, September 24, 2012. <https://www.telegraph.co.uk/news/uknews/defence/9552547/The-air-force-men-who-fly-drones-in-Afghanistan-by-remote-control.html>.

Blair, Dave and Karen House. "Avengers with Wrath: Moral Agency and Trauma Prevention for Remote Warriors." *Lawfare*, November 12, 2017. <https://www.lawfareblog.com/avengers-wrath-moral-agency-and-trauma-prevention-remote-warriors>.

Boffey, Philip M. "Engineer who Opposed Launching Challenger Sues Thiokol for \$1 Billion." *New York Times*, January 29, 1987. <http://www.nytimes.com/1987/01/29/us/engineer-who-opposed-launching-challenger-sues-thiokol-for-1-billion.html>.

Boyd, John. "The Essence of Winning and Losing." The Internet Archive, accessed May 17, 2018. <https://web.archive.org/web/20110324054054/http://www.danford.net/boyd/essence.htm>.

Bureau of Labor Statistics. "Occupational Employment Statistics." United States Department of Labor. Accessed May 17, 2018. <https://www.bls.gov/oes/current/oes432021.htm>.

Camden, Jim. "Air Force Reprimands, Fines Colonel Officer who was in Charge of Flight Operations, Says B-52 Crash Will Haunt Him for Rest of his Life." *Spokane Spokesman-*

- Review*, May 23, 1995. <http://www.spokesman.com/stories/1995/may/23/air-force-reprimands-fines-colonel-officer-who/>
- Cantwell, Houston and Alfred Rosales. "RPA Lost Link: What do we do now?" *Combat Edge* 26 no. 1 (2017).
- Clausewitz, Carl von. *On War*. Translated and edited by Michael Howard and Peter Paret. Princeton: Princeton University Press, 1976.
- Coldewey, Devin. "Uber Vehicle Reportedly Saw but Ignored Woman it Struck." *Techcrunch*, May 7, 2018. <https://techcrunch.com/2018/05/07/uber-vehicle-reportedly-saw-but-ignored-woman-it-struck/>.
- Crane, Conrad C. "The Lure of the Strike." *Parameters: The U.S. Army War College Quarterly* 43, no. 2 (2013).
http://www.strategicstudiesinstitute.army.mil/pubs/parameters/issues/Summer_2013/1_Crane_SpecialCommentary.pdf
- Cutler, Joyce E. "Federal Judge Dismisses Boisjoly's Lawsuits Against Thiokol." *Deseret News*, September 2, 1988. <https://www.deseretnews.com/article/15938/FEDERAL-JUDGE-DISMISSES-BOISJOLYS-LAWSUITS-AGAINST-THIOKOL.html>
- Danaher, John. "Robots, Law and the Retribution Gap." *Ethics and Information Technology* 18, no. 4 (2015): 299-309.
- Diehl, Alan E. *Silent Knights: Blowing the Whistle on Military Accidents and Their Cover-Ups*. Washington, DC: Brassey's, 2002.
- Dillow, Clay. "Pentagon Report: The F-35 is Still a Mess." *Fortune*, March 10, 2016. <http://fortune.com/2016/03/10/the-f-35-is-still-a-mess/>.
- Dockterman, Eliana. "Robot Kills Man at Volkswagen Plant." *TIME Magazine*, January 2, 2015. <http://time.com/3944181/robot-kills-man-volkswagen-plant/>.
- Drew, James. "USAF to Automate MQ-9 Takeoffs and Landings." *Flight Global*, May 4, 2016. <https://www.flightglobal.com/news/articles/usaf-to-automate-mq-9-takeoffs-and-landings-424975/>.
- Federal Reserve Bank of New York. "Automated Clearing Houses (ACHs)." Accessed May 17, 2018. <https://www.newyorkfed.org/aboutthefed/fedpoint/fed31.html>.
- Fraley, Jill M. "Re-Examining Acts of God." *Pace Environmental Law Review* 27, no. 3 (2010): 669-690.
- Gadher, Dipesh and Toby Harnden. "Islamic State Hackers Publish Hit List of US Drone Pilots." *The Weekend Australian*, May 2, 2016. <https://www.theaustralian.com.au/news/world/islamic-state-hackers-publish-hit-list-of-us-drone-pilots/news-story/3b64879379c008be78f07000deacc4f1>.

- Galison, Peter. "An Accident of History." in *Atmospheric Flight in the Twentieth Century*, ed. Peter Galison and Alex Roland (London: Kluwer Academic Publishers, 2000), 3-4.
- Garfinkel, Simson. "History's Worst Software Bugs." *Wired*, November 8, 2005. <https://www.wired.com/2005/11/historys-worst-software-bugs/>.
- Gates, David. *Sky Wars*. London: Reaktion Books, 2003.
- Gershgorn, Dave. "The Case Against Understanding Why AI Makes Decisions," *Quartz*, January 31, 2018. <https://qz.com/1192977/the-case-against-understanding-why-ai-makes-decisions/>.
- Gilboa, Eytan. "The CNN Effect: The Search for a Communication Theory of International Relations." *Political Communication* 22, no. 1 (2005): 27-44.
- Gladstone, Rick. "If U.S. Attacks North Korea First, Is That Self-Defense?" *New York Times*, August 10, 2017. <https://www.nytimes.com/2017/08/10/world/asia/us-north-korea-preemptive-attack-questions-answers.html>.
- Gless, Sabine, Emily Silverman, and Thomas Weigend. "If Robots Cause Harm, Who Is to Blame: Self-Driving Cars and Criminal Liability." *New Criminal Law Review* 19, no. 3 (2016): 412-436.
- Golson, Jordan and Dieter Bohn. "All new Tesla Cars Now Have Hardware for 'Full Self-driving Capabilities'." *The Verge*, October 19, 2016. <https://www.theverge.com/2016/10/19/13340938/tesla-autopilot-update-model-3-elon-musk-update>.
- Gordon, Michael R. and Bernard E. Trainor. *The Generals' War: The Inside Story of the Conflict in the Gulf*. Boston: Little, Brown and Company, 1995.
- Grossman, Dave. *On Killing*. Boston: Little, Brown and Company, 1995.
- Guttman, John. "Norden M-1 Bombsight." *Military History* 25, no. 5 (2008): 25. <https://search-proquest-com.usnwc.idm.oclc.org/docview/212664512>.
- Guy Norris, "Auto-GCAS Saves Unconscious F-16 Pilot—Declassified USAF Footage," *Aviation Week*, September 13, 2016, <http://aviationweek.com/air-combat-safety/auto-gcas-saves-unconscious-f-16-pilot-declassified-usaf-footage>.
- Harris, Arthur. *Bomber Offensive*. New York: The MacMillan Company, 1947.
- Himes, Kenneth R. *Drones and the Ethics of Targeted Killing*. Lanham: Rowman & Littlefield, 2016.
- Hobbs, Alan. "Human Factors of Remotely Piloted Aircraft Systems: Lessons from Incident Reports." National Aeronautics and Space Administration. Last modified February 10,

2017. <https://www.nasa.gov/mediacast/human-factors-of-remotely-piloted-aircraft-systems-lessons-from-incident-reports>.
- Hughes, Wayne P. *Fleet Tactics and Coastal Combat*. Annapolis: Naval Institute Press, 2000.
- Huntington, Samuel P. *The Clash of Civilizations and the Remaking of World Order*. New York: Touchstone, 1997.
- International Business Times India*. "Robot Kills Man at Gurgaon Factory." August 13, 2015. <http://www.ibtimes.co.in/robot-kills-man-gurgaon-factory-642723>.
- International Organization for Standardization. "ISO 13482:2014: Robots and Robotic Devices – Safety Requirements for Personal Care Robots." Accessed May 1, 2018. <https://www.iso.org/standard/53820.html>.
- "Internet History: From ARPANet to Broadband," *The Congressional Digest* 86 no. 2 (2007). <http://congressionaldigest.com/issue/network-neutrality/internet-history/>.
- Jansen, Bart. "Amtrak Train Derailed on Tracks that had Automatic-braking Technology—But it was Still Being Tested." *USA Today*, December 19, 2017, <https://www.usatoday.com/story/news/2017/12/19/speeding-amtrak-train-derailed-track-without-automatic-braking/964483001/>.
- Jones, Meg Leta. "The Ironies of Automation Law: Tying Policy Knots with Fair Automation Practices Principles." *Vanderbilt Journal of Entertainment and Technology Law* 18, no. 1 (2015): 77-134.
- Kant, Immanuel. *Foundations of the Metaphysics of Morals and Critical Essays*. Edited by Robert Paul Wolff, translated by Lewis White Beck. London: MacMillan Pub Co, 1990.
- Kean, Thomas H., Lee Hamilton, and National Commission on Terrorist Attacks upon the United States. *The 9/11 Commission Report: Final Report of the National Commission on Terrorist Attacks upon the United States*. Washington, DC: National Commission on Terrorist Attacks upon the United States, 2004.
- Khoury, Nabeel. "In Yemen, Drones Aren't a Policy." *Cairo Review of Global Affairs*, October 23, 2013. <https://www.thecaireview.com/tahrir-forum/in-yemen-drones-arent-a-policy/>.
- Knight, Will. "The Dark Secret at the Heart of AI." *MIT Technology Review*, April 11, 2017. <https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/>.
- Kolata, Gina. "How Fen-Phen, a Diet 'Miracle,' Rose and Fell." *New York Times*, Sept 23, 1997. <http://www.nytimes.com/1997/09/23/science/how-fen-phen-a-diet-miracle-rose-and-fell.html>.
- Kranzberg, Melvin. "Kranzberg's Laws." *Technology and Culture* 27, no. 3 (1986): 544-560. <http://www.jstor.org/stable/3105385>.

- Laris, Michael and Faiz Siddiqui. "Video Shows Autonomous Uber and Backup Driver Failing to Protect Pedestrian." *Washington Post*, March 22, 2018.
<https://www.washingtonpost.com/news/dr-gridlock/wp/2018/03/22/video-shows-autonomous-uber-and-backup-driver-failing-to-protect-pedestrian/>.
- Launius, Roger D. "Sputnik 'Was Force for World Peace'." *BBC News*, October 4, 2007.
<http://news.bbc.co.uk/2/hi/science/nature/6965717.stm>.
- Lee, Timothy B. "Report: Software Bug Led to Death in Uber's Self-driving Crash." *Ars Technica*, May 7, 2018. <https://arstechnica.com/tech-policy/2018/05/report-software-bug-led-to-death-in-ubers-self-driving-crash/>.
- Lindblom, Mike and David Gutman. "NTSB Report: Amtrak Engineer Missed Speed-limit Signs Before Train Crashed South of Tacoma." *Seattle Times*, January 25, 2018.
<https://www.seattletimes.com/seattle-news/transportation/ntsb-report-amtrak-engineer-missed-speed-limit-sign-before-the-train-crashed-on-a-curve-south-of-tacoma/>.
- Lockheed Martin. F-35 Software Development. Last accessed Feb 11, 2018,
<https://www.f35.com/about/life-cycle/software>.
- Lohmann, Melinda Florina. "Liability Issues Concerning Self-Driving Vehicles." *European Journal of Risk Regulation* 7, no. 2 (2016): 335-340.
- Lombardo, Cara. "Uber CEO Says Self-Driving Cars Are 'Student Drivers'." *Wall Street Journal*, April 12, 2018. <https://www.wsj.com/articles/uber-ceo-says-self-driving-cars-are-student-drivers-1523538431>.
- Marlantes, Karl. *What it is Like to Go to War*. New York: Atlantic Monthly Press, 2011.
- Marshour, George, Stuart Forman, and Jason Campagna. "Mechanisms of General Anesthesia: From Molecules to Mind." *Best Practice & Research Clinical Anaesthesiology* 19, no. 3 (2005): 349-364. <https://www.sciencedirect.com/science/article/pii/S1521689605000054>.
- Maslow, Abraham H. *The Psychology of Science; A Reconnaissance*. New York: Harper & Row, 1966.
- Massachusetts Institute of Technology. "Moral Machine." Last accessed February 11, 2018,
<http://moralmachine.mit.edu>.
- Matsuzaki, Hironori and Gesa Lindemann. "The Autonomy-Safety-Paradox of Service Robotics in Europe and Japan: A Comparative Analysis." *AI & Society* 31, no. 4 (2016): 501-517.
- Mattis, James. "Press Gaggle by Secretary Mattis En Route to Washington, D.C." U.S. Department of Defense, February 17, 2018.
<https://www.defense.gov/News/Transcripts/Transcript-View/Article/1444921/press-gaggle-by-secretary-mattis-en-route-to-washington-dc/>.

- Military.com*. "Navy F-35C Landed So Precisely, it Tore up a Runway." August 18, 2016. <https://www.military.com/dodbuzz/2016/08/18/navy-f-35c-landed-so-precisely-it-tore-up-a-runway>.
- Mouloua, Mustapha, Richard Gilson, Eleni Daskarolis-Kring, Jason Kring, and Peter Hancock. "Ergonomics of UAV/UCAV Mission Success: Considerations for Data Link, Control, and Display Issues." *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting*, Minneapolis/St. Paul, MN, 2001. Santa Monica: Human Factors and Ergonomics Society, 2001. https://www.researchgate.net/profile/Peter_Hancock2/publication/238075274_Ergonomics_of_UAVUCAV_Mission_Success_Considerations_for_Data_Link_Control_and_Display_Issues/links/00b7d51c8604b9173a000000/Ergonomics-of-UAV-UCAV-Mission-Success-Considerations-for-Data-Link-Control-and-Display-Issues.pdf.
- Nambu, Tomoko. "Legal Regulations and Public Policies for Next-Generation Robots in Japan." *AI & Society* 31, no. 4 (2016): 483-500.
- National Highway Traffic Safety Administration. "NHTSA Announces Final Rule Requiring Rear Visibility Technology." March 31, 2014. <https://www.nhtsa.gov/press-releases/nhtsa-announces-final-rule-requiring-rear-visibility-technology>.
- National Transportation Safety Board. *Loss of Thrust in Both Engines After Encountering a Flock of Birds and Subsequent Ditching on the Hudson River, US Airways Flight 1549, Airbus A320-214, N106US, Weehawken, New Jersey, January 15, 2009*. May 4, 2010. <https://www.nts.gov/investigations/AccidentReports/Reports/AAR1003.pdf>.
- National Transportation Safety Board. *Preliminary Report: Highway HWY18MH010*. Accessed May 29, 2018. <https://www.nts.gov/investigations/AccidentReports/Reports/HWY18MH010-prelim.pdf>.
- Nazeri, Zohreh. "Cross-Database Analysis to Identify Relationships Between Aircraft Accidents and Incidents." PhD diss., George Mason University, 2007. ProQuest (AAT 3285768).
- Obama, Barack, U.S. President. Address, National Defense University, Fort McNair. Washington, DC, May 23, 2013. <https://obamawhitehouse.archives.gov/the-press-office/2013/05/23/remarks-president-national-defense-university>.
- Online Community Library Center, "OCLC Prints Last Library Catalog Cards," October 1, 2015, <https://www.oclc.org/en/news/releases/2015/201529dublin.html>.
- Overy, Richard. *The Bombing War: Europe 1939-1945*. London: Penguin Press, 2013.
- O'Mara, Raymond P. "The Socio-Technical Construction of Precision Bombing: A Study of Shared Control and Cognition by Humans, Machines, and Doctrine During World War II." PhD diss., Massachusetts Institute of Technology, 2011. <http://hdl.handle.net/1721.1/67754>.

- Perrin, Chad. "The Danger of Complexity: More Code, More Bugs," *Tech Republic*, February 1, 2010. <https://www.techrepublic.com/blog/it-security/the-danger-of-complexity-more-code-more-bugs/>.
- Pogue, Forrest C. "It's Hard to Decide Whether Patton Was Drunk From War," Review of *Patton: Ordeal and Triumph*, by Ivan Obolensky, *The Washington Post* (December 18, 1964): A18. ProQuest, <https://search-proquest-com.usnwc.idm.oclc.org/docview/142086667?pq-origsite=summon>.
- President's Commission on the Space Shuttle Challenger Accident. *Report to the President by the Presidential Commission on the Space Shuttle Challenger Accident*. Washington, DC, 1986.
- Rimer, Sara. "Once a Friendly Fixture, a Telephone Operator Finds Herself Obsolete." *New York Times*, June 4, 1996. <https://www.nytimes.com/1996/06/04/us/once-a-friendly-fixture-a-telephone-operator-finds-herself-obsolete.html>.
- Russell, Stuart et al. "Autonomous Weapons: An Open Letter from AI & Robotics Researchers." Future of Life Institute. Accessed April 13, 2018. <https://futureoflife.org/open-letter-autonomous-weapons/>.
- SAE International. "Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems." Accessed May 1, 2018. https://www.sae.org/standards/content/j3016_201401/.
- Shah, Aqil. "Drone Blowback in Pakistan is a Myth. Here's Why." *Washington Post*, May 17, 2016. <https://www.washingtonpost.com/news/monkey-cage/wp/2016/05/17/drone-blow-back-in-pakistan-is-a-myth-heres-why/>.
- Shead, Sam. "Amazon Now Has 45,000 Robots in its Warehouses." *Business Insider*, January 3, 2017. <http://www.businessinsider.com/amazons-robot-army-has-grown-by-50-2017-1>.
- Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/consequentialism/>.
- Statt, Nick. "iPhone Manufacturer Foxconn Plans to Replace Almost Every Human Worker with Robots." *The Verge*, December 30, 2016. <https://www.theverge.com/2016/12/30/14128870/foxconn-robots-automation-apple-iphone-china-manufacturing>.
- Stimson Center. "Recommendations and Report on the Task Force on Drone Policy." Co-chairs, John Abizaid and Rosa Brooks; project director, Rachel Stohl. Washington, DC: Stimson Center, 2014. Quoted in Kenneth R. Himes. *Drones and the Ethics of Targeted Killing*. Lanham: Rowman & Littlefield, 2016.
- TASS Russian News Agency. "Artificial Intelligence to Replace Pilot in Aircraft Cockpit – Russian Senator." November 1, 2017. <http://tass.com/defense/973707>.
- Tetteh, Edem G. "Human Factors Analysis of Commercial Aircraft Accidents in the United States: 1960-2000." *IISE Annual Conference Proceedings* (2006): 1-7.

- United Kingdom Department of Transport. *The Pathway to Driverless Cars: Summary Report and Action Plan*. London, England, 2015.
- U.S. Census Bureau. “Quarterly E-Commerce Sales, 1st Quarter 2018.” Accessed May 17, 2018. https://www.census.gov/retail/mrts/www/data/pdf/ec_current.pdf.
- U.S. Food and Drug Administration. “The Drug Development Process.” Last accessed February 11, 2018. <https://www.fda.gov/ForPatients/Approvals/Drugs/ucm405622.htm#Approval>.
- U.S. House of Representatives. *Investigation of the Challenger Accident*. Committee on Science and Technology. Washington, DC: October 29, 1986.
- U.S. Navy. “U.S. Navy Fact File - Tomahawk Cruise Missile.” Last modified April 10, 2017. http://www.navy.mil/navydata/fact_display.asp?cid=2200&tid=1300&ct=2.
- U.S. Occupational Safety and Health Administration. Accident Database. Accessed January 30, 2018. https://www.osha.gov/pls/imis/AccidentSearch.search?acc_keyword=%22Robot%22&keyword_list=on.
- Wall Street Journal*, “Automation and Liability,” December 12, 1961, 14.
- Washington Post*. “Colonel Pleads Guilty to Dereliction in Crash.” May 20, 1995. https://www.washingtonpost.com/archive/politics/1995/05/20/colonel-pleads-guilty-to-dereliction-in-crash/8179f31b-5353-4dc8-9403-d6ea99005e42/?utm_term=.9ec2bb52fb9c.
- Wiener, Norbert. *The Human Use of Human Beings: Cybernetics and Society*. New York: Avon Books, 1954. Quoted in Meg Leta Jones. “The Ironies of Automation Law: Tying Policy Knots with Fair Automation Practices Principles.” *Vanderbilt Journal of Entertainment and Technology Law* 18, no. 1 (2015): 77-134.
- WOOD TV. “Fine issued in worker’s death at Ionia plant.” February 2, 2016. <http://woodtv.com/2016/02/02/fine-issued-in-workers-death-at-ionia-plant/>.
- World Bank. “U.S. Fixed Telephone Subscriptions (per 100 People).” World Bank Open Data. Accessed May 17, 2018. <https://data.worldbank.org/indicator/IT.MLT.MAIN.P2?end=2016&locations=US&start=1960>.